

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 19, 2002, 02:33:05 ; Search time 200 Seconds
(without alignments)
2438.019 Million cell updates/sec

Title: US-09-846-808-1
Perfect score: 1510
Sequence: 1 MSKRSNRKFWLWMLILFTP.....ALAMLSIGYVGGSGIKFIL 284

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=framet_p2n model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US09846808/runat_18102002_140434_26005/app_query.fasta_1.455
-DB=N_Geneseq_032802 -QMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS-human4.0.cdi
-LIST=45 -DOCALL=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09846808 @CGN_1.1.48 @runat_18102002_140434_26005 -NCPU=6 -ICPU=3
-NO_XLPYX -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	517	34.2	858	20	AAX34757	DNA encoding OMP-1
2	481.5	31.9	852	20	AAX34755	DNA encoding OMP-1
3	466	30.9	882	20	AAX34766	DNA encoding P30-6
4	455.5	30.2	900	20	AAX34768	DNA encoding P30-8
5	451	29.9	828	20	AAX34736	DNA encoding OMP-1
6	448	29.7	882	20	AAX34765	DNA encoding P30-5
7	446	29.5	894	20	AAX34749	DNA encoding OMP-1
8	432	28.6	840	20	AAX34754	DNA encoding OMP-1
9	431.5	28.6	852	20	AAX34744	DNA encoding OMP-1
10	415	27.5	891	20	AAX34767	DNA encoding P30-7
11	414.5	27.5	888	20	AAX34753	DNA encoding OMP-1
12	411.5	27.3	843	21	AAC68705	Ehrlichia chafeen
13	411.5	27.3	4683	19	AAV07179	Ehrlichia chafeen
14	411.5	27.3	4683	21	AAC68716	Ehrlichia chafeen
15	411.5	27.3	4683	22	AAV07578	DNA encoding varia
16	408.5	27.1	843	20	AAX34748	DNA encoding OMP-1
17	389	25.8	840	21	AAD01294	Ehrlichia canis im
18	389	25.8	2037	21	AAD01295	Ehrlichia canis im
19	387.5	25.7	837	20	AAX34747	DNA encoding OMP-1
20	387.5	25.7	837	21	AAC68704	Ehrlichia chafeen
21	386	25.6	843	20	AAX34762	DNA encoding P30-2
22	386	25.6	867	19	AAX34759	DNA encoding P30 p
23	383.5	25.4	864	19	AAV07176	Cowdria ruminantlu
24	383.5	25.4	864	21	AAC68699	DNA encoding major
25	383.5	25.4	864	22	AAV07575	DNA encoding OMP-1
26	380	25.2	861	20	AAX34746	Ehrlichia chafeen
27	380	25.2	861	21	AAC68703	Ehrlichia chafeen
28	379.5	25.1	845	20	AAX34743	DNA encoding OMP-1
29	377	25.0	817	20	AAX34758	DNA encoding P28 p
30	375.5	24.9	756	20	AAX34742	Ehrlichia chafeen
31	371.5	24.6	830	21	AAC68706	Ehrlichia chafeen
32	371.5	24.6	843	20	AAX34745	DNA encoding OMP-1
33	370	24.5	842	19	AAV07177	Ehrlichia chafeen
34	370	24.5	842	21	AAC68700	Ehrlichia chafeen
35	370	24.5	842	22	AAV07576	DNA encoding major
36	363	24.0	924	20	AAX34761	DNA encoding P30-1
37	363	24.0	1607	21	AAD01292	Ehrlichia canis im
38	342.5	22.7	1570	19	AAV07180	Ehrlichia canis VS
39	342.5	22.7	1570	22	AAV07583	DNA encoding varia
40	341.5	22.6	831	20	AAX34764	DNA encoding P30-4
41	341.5	22.6	864	21	AAC68707	Ehrlichia canis VS
42	339.5	22.5	1570	21	AAC68717	Ehrlichia canis 28
43	335.5	22.1	864	20	AAX34760	DNA encoding P30a
44	335.5	21.6	852	20	AAX34763	DNA encoding P30-3
45	324.5	21.5	849	21	AAD01293	Ehrlichia canis im

ALIGNMENTS

RESULT 1
AAX34757
ID AAX34757 standard; DNA; 858 BP.
XX
AC AAX34757;
XX
DT 05-JUL-1999 (first entry)
XX
DE DNA encoding OMP-1Y protein.
XX
KW Outer membrane protein; OMP; Ehrlichia chafeensis; E. canis; P30;
KW detection; dog; ss.
XX
OS Ehrlichia chafeensis.
XX
PN WO9913720-A1.
XX
PD 25-MAR-1999.

```

XX 18-SEP-1998; 98WO-US19600.
XX
XX 19-SEP-1997; 97US-0059353.
XX
XX (OHIS ) UNIV OHIO STATE.
XX
XX Ohashi N, Rikihisa Y;
XX
XX WPI; 1999-254290/21.
XX P-PSDB; AAY06957.
XX
XX Novel outer membrane proteins from Ehrlichia chaffeensis and
XX Ehrlichia canis
XX
XX Disclosure; Fig 17A; 55pp; English.
XX
XX The invention provides isolated outer membrane proteins (OMP) from
XX Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
XX of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
XX in AAY06943-958. The E. canis proteins form part of the P30 family and
XX consist of proteins shown in AAY06959-970. The proteins and genes are
XX used to detect E. chaffeensis in patients and E. canis in dogs.
XX
XX Sequence 858 BP; 293 A; 135 C; 158 G; 272 T; 0 other;

Alignment Scores:
Pred. No.: 8,84e-49 Length: 858
Score: 517.00 Matches: 109
Percent Similarity: 59.59% Conservative: 65
Best Local Similarity: 37.33% Mismatches: 96
Query Match: 34.24% Indels: 22
DB: 20 Gaps: 8

US-09-846-808-1 (1-284) x AAX34757 (1-858)

QY 5 SerAsnArgLys-----PheValLeuTrpValMetLeu-----ileLeuPheThr 19
Db 4 AATAATAGAAAAAGTTTATATAGAGTGCACTACTACGAAGCTTATTATTATTCACA 63
QY 20 ProHisIleSerLeuAlaSerValLeuAsnAspHisAsn-----SerMetTyr 35
Db 64 TCTGAGCCCTCTTCTACAGAAATGTAAGTAACCATCTATTATTTAAACCTAGGTTATAT 123
QY 36 ValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSer 55
Db 124 ATCAGTGGACAATATAGACAGGAGTTTCTCATTTTAGCAAAATTTTCAGTCAAGAAACC 183
QY 56 AlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThr 75
Db 184 AACTACAATACTACTCAACTAGTTGGCTTAAAAAGGACATCATGTCATAGGGAACAGT 243
QY 76 GlyIleLysAspAsnThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArg 95
Db 244 AATATCACACCTACACAAATTTCAACTTTCCTTACATTGCAGAAATTTCAAGACAATGCC 303
QY 96 LeuGlyPheSerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGlu 115
Db 304 ATAGTTTCAGTGGGCAATTTGATCTACTGTATTCGAGAAATTTAGAAATGAGTAGAG 363
QY 116 LeuSerTyrGluThrPheHisIleLysAsnAsnGlyTyrLysArgIleAspCysGluLys 135
Db 364 GCTTCTTATGAACAATTGATGTTAAAAATCCAGAGGAGTCTCTACAGCGCATACAGG 423
QY 136 HisPheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsn----- 153
Db 424 TATTTTGCATAGCAGCTGCTATGAGTGGCACTAAATAAATCTAGTCTCTGATGACACAAGA 483
QY 154 LysTyrValThrLeuLeuAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCys 173
Db 484 AATTCACCTCATGAGAAATACGGGGTATCAATTTTCATCAGTAATGATAAATGGGTGT 543
QY 174 TyrAsp-----ValAspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPhe 191

```

```

Db 544 TACAATTTTACATTAGATATATACCA-----GTAGTACCGTATGTATGCGCAGGAATA 597
QY 192 GlyValAspThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLys 211
Db 598 GGAGGAGATTTTCATAGAGTTTTTAATGATGATTACATGTTAGTTTGTCTATCAAGCAAG 657
QY 212 LeuGlyAlaSerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHis 231
Db 658 GTAGGTATTAGTTATTCTATATCCCTGAAGTAAGTTTATTCTTAACGGATATTACCAT 717
QY 232 GlyLeuPheGlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSer 251
Db 718 AAAGTAACAGGTAACAGATTTAAAAAAGTTACACGTTCAACACGTAAGTGTATTTA---AGT 774
QY 252 ProThrProProAsnSerLysProHisValHisThrAlaLeuAlaMetLeuSerIle 271
Db 775 GAGCCTCCTAAG-----TTTCACATCTCGAGTTGTACACTCAATGTT 816
QY 272 GlyTyrTyrGlyGlySerIleGlyIleLysPheIle 283
Db 817 GGGTACTTTGGTGGCGAAATTTGGAGTAAGATTTATA 852

RESULT 2
AAX34755
ID AAX34755 standard; DNA; 852 BP.
XX
XX AAX34755;
XX
XX 05-JUL-1999 (first entry)
XX
XX DNA encoding OMP-1W protein.
XX
XX Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
XX detection; dog; ss.
XX
XX Ehrlichia chaffeensis.
XX
XX WO9913720-A1.
XX
XX 25-MAR-1999.
XX
XX 18-SEP-1998; 98WO-US19600.
XX
XX 19-SEP-1997; 97US-0059353.
XX
XX (OHIS ) UNIV OHIO STATE.
XX
XX Ohashi N, Rikihisa Y;
XX
XX WPI; 1999-254290/21.
XX P-PSDB; AAY06955.
XX
XX Novel outer membrane proteins from Ehrlichia chaffeensis and
XX Ehrlichia canis
XX
XX Disclosure; Fig 15A; 55pp; English.
XX
XX The invention provides isolated outer membrane proteins (OMP) from
XX Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
XX of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
XX in AAY06943-958. The E. canis proteins form part of the P30 family and
XX consist of proteins shown in AAY06959-970. The proteins and genes are
XX used to detect E. chaffeensis in patients and E. canis in dogs.
XX
XX Sequence 852 BP; 295 A; 115 C; 141 G; 301 T; 0 other;

Alignment Scores:
Pred. No.: 9,01e-45 Length: 852
Score: 481.50 Matches: 106
Percent Similarity: 55.60% Conservative: 48
Best Local Similarity: 38.27% Mismatches: 108
Query Match: 31.89% Indels: 15

```

DB: 20 Gaps: 5

US-09-846-808-1 (1-284) x AAX34755 (1-852)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
||| |||::: :::: ||| ||||||: ::::: ||| ::::: |||

Db 43 GTATGTTAGTGTCACTACTACCTAAATCTTGTCACAACCTTAATAATTATAAT 102
::: |||::: |||::: |||::: |||::: |||::: |||::: |||

QY 33 Ser-----MetTyrrValGlyIleLeuTyrrLysProAlaArgGlnHisLeu 47
::: |||::: |||::: |||::: |||::: |||::: |||::: |||

Db 103 AACACTAAGTGCACCTGGCGTATATGTCAGTGGACAATATAAACCTACTGTTCTCACCATT 162
::: |||::: |||::: |||::: |||::: |||::: |||::: |||

QY 48 SerLysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLys 67
||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||

Db 163 AGTAATTTTTCTACTTAAGAAGAACTATATCATGACACTAAAAGCTTATATGAGCATGACAAA 222
::: |||::: |||::: |||::: |||::: |||::: |||::: |||

QY 68 AspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIleLysTyrr 87
|||::: ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||

Db 223 GATATTAAAGTCT-----ATTACAGATATAACAACAATAAAAAATTCACAATTCCTTAT 276
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||

QY 88 AsnProTyrrTyrrGluAsnAsnArgLeuGlyPheSerGlyIlePheGlyTyrrTyrrAsn 107
||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||

Db 277 AACACAAAATTTCAAGATAATGCTGTAGCTTCAGTGCACGCTGTTGGATATATATCCCAA 336
||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||

QY 108 LysAsnPheArgIleGluSerGluLeuSerTyrrGluThrPheHisIleLysAsnAsnGly 127
::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||

Db 337 GACAGTCCAAGGGTTGAGGTAGAAATGCTTCTATGAAGAATTTACGCTTAAAAATCCTGGT 396
::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||

QY 128 --TyrrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGly 146
||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||

Db 397 AYTTCGTAGTAGTGAAGCCTTCAGGTATATGCTTTAGCAAGGAATTCGATAAATCTT 456
||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||

QY 147 SerAsnAsnProAlaAsnAsnLysTyrrValThrLeuIleAsnAsnGlyIleSerLeuThr 166
||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||

Db 457 CAAAAATATCCTGAACAAATAAGTATGTTTATAAACAACAATGGCTTATCTGTGCGA 516
||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||

QY 167 SerAlaLeuIleAsnValCystyrAspValAspGlyLeuLysHisAsnIleIleThrTyrr 186
||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||

Db 517 TCCATTATAATCAATGGCTGTTATGATTTTTCTTTAAACAATTTAAAGATATCACCTTAC 576
||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||

QY 187 SerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrrThrThrLysPhe 206
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||

Db 577 ATATGCGTAGGGTTGGTGGGACATATAGAAATTTTATAGTGTGAAGTTTAAATTT 636
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||

QY 207 SerTyrrGlnGlyLysLeuGlyAlaSerTyrrThrValSerProGlnValSerValPheIle 226
::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||

Db 637 GCTTATCAAGTAAAGTAGGTATCAGTTATCCATTATCTCTAATATGATTATATTTGCT 696
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||

QY 227 GluGlyTyrrTyrrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsnTyrrPro 246
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||

Db 697 GAGGGATATTACCNTAAGGTACAGGAAATAAATTTAAACAATTTAAATGTTCAACAGCTT 756
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||

QY 247 CysAspTyrrProserProthrProProAsnSerLysProHisValHisThrThrAlaLeu 266
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||

Db 757 GTTAGTCTT-----ACAGCTATPCCTAAG---TCTACTTTTCGAGTA 795
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||

QY 267 AlaMetLeuSerIleGlyTyrrTyrrGlyGlySerIleGlyIleLysPheIle 283
||| |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||

Db 796 GCTACTCTTAATGTCAGTATTTTCGGTAGTGAATTTGGGTTAAATTTATA 846
|||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||

RESULT 3

AAX34766

ID AAX34766 standard; DNA; 882 BP.

XX AAX34766;

AC AAX34766;

XX XX

DT DT

DE DE

XX XX

KW Outer membrane protein; OMP; Ehrlichia chafeensis; E. canis; P30;

KX detection; dog; ss.

XX XX

AC AAX34756;
 XX 05-JUL-1999 (first entry)
 XX DNA encoding OMP-1X protein.
 XX Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 KW detection; dog; ss.
 XX Ehrlichia chaffeensis.
 OS
 XX
 XX W09913720-A1.
 PN
 XX
 XX 25-MAR-1999.
 PD
 XX 18-SEP-1998; 98WO-US19600.
 PF
 XX 19-SEP-1997; 97US-0059353.
 PR
 XX (OHIS) UNIV OHIO STATE.
 PA
 XX Ohashi N, Rikihisa Y;
 PI
 XX WPI; 1999-254290/21.
 DR
 XX P-PSDB; AAY06956.
 PT
 XX Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX
 XX Disclosure; Fig 16A; 55pp; English.
 PS
 XX The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX
 XX Sequence 828 BP; 281 A; 131 C; 136 G; 280 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 2.43e-41 Length: 828
 Score: 451.00 Matches: 101
 Percent Similarity: 53.76% Conservative: 49
 Best Local Similarity: 36.20% Mismatches: 99
 Query Match: 29.87% Indels: 30
 DB: 20 Gaps: 6
 US-09-846-808-1 (1-284) x AAX34756 (1-828)
 QY 14 MetLeuLeuLeuPheThrProHisLeuSerLeuAlaSerValLeuAsnAspHis----- 31
 DB 37 CTATTATCATATGTTGTTACTTAACATATCTTTCCAGAAACATTAACCAATTAACACTGAT 96
 QY 32 -----AsnSerMetTyrValGlyLeuGlnTyrLysProAlaArgGlnHisLeuSerLys 49
 DB 97 AACCTTCTGGGTATATATAAGTGGCAATATAACACAGGATTTCTCATTTTCACCAA 156
 QY 50 LeuLeuLeuLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeu 69
 DB 157 TTTTCAGTCAAGAAATCATATGATAAATCAATTAACCAATTAACCAATTAACCAATTAAC 216
 QY 70 LeuAsnAspLeuThrGlyLysAspAsnThrAsnPheAsnIleLysLysTyrAsnPro 89
 DB 217 ATTTCTACTAGTACCTTAAATATT-----AATACAGATTTTAATATCCCTATAAGTA 270
 QY 90 TyrTyrGluAsnArgLeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsnLysAsn 109
 DB 271 ACATTTCAAAATACATTAACAGCTTTAGTGGAGCTATTGGTATTCTCATCCACAGGG 330
 QY 110 PheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsnGlyTyrLys 129
 DB 331 GCAAGATTTGAGCTTGAAGGTTCTTATGAAGAATTGATGTGACAGATCCTGGAGACTGC 390

QY 130 ArgIle---AspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGlySerAsn 148
 DB 391 TTAATAAAGATACCTATAGATATTTTCGTTTAGTAGAACCACATCAGGTTCTAGC--- 447
 QY 149 AsnProAlaAsnLysTyrValThrLeuLeuAsnGlyIleSerLeuThrSerAla 168
 DB 448 ---CCTACCTCAACAACATATACTGTTATGAGAAATGATGGTTCATCTCTGTT 504
 QY 169 LeuLeuAsnValCysTyrAspValAspGlyLeuLysHisAsnIleIleThrTyrSerCys 188
 DB 505 ATATTTAATGGCTGTTATGACATCTTTTAAAGGATTTAGAGATATACCTTTATGATGT 564
 QY 189 LeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyr 208
 DB 565 GTTGGTGTAGTGGAGATTTTATAGAATTTTTCGACCATTAACATTAATTAGCATAC 624
 QY 209 GlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIleGluGly 228
 DB 625 CAAGGCAAGTTAGGTATCAATTATCACTTATCGACTCAAGCAAGCGTATTATTGATGGA 684
 QY 229 TyrTyrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsnTyrProCysAsp 248
 DB 685 TATTATCATAGGTTATAGGAATCAATTCACAATCTAAATGTTCAA----- 732
 QY 249 TyrProSerProThrProProAsnSerLysProHisValHisThrThr----- 264
 DB 733 -----CACGTGGCTAGTACAGATTTTGGACCT 759
 QY 265 -----AlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyIleLys 281
 DB 760 GTATACGACAGTACCCACTTAACATTTGTTATTTGGTGGTGAATCGAATTAGA 816
 RESULT 6
 AAX34765
 ID AAX34765 standard; DNA; 882 BP.
 XX
 XX AAX34765;
 XX
 DT 05-JUL-1999 (first entry)
 XX DNA encoding P30-5 protein.
 DE
 XX Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 KW detection; dog; ss.
 XX
 XX Ehrlichia canis.
 OS
 XX
 XX W09913720-A1.
 PN
 XX
 PD 25-MAR-1999.
 XX
 PF 18-SEP-1998; 98WO-US19600.
 PR
 XX 19-SEP-1997; 97US-0059353.
 XX (OHIS) UNIV OHIO STATE.
 PA
 XX Ohashi N, Rikihisa Y;
 PI
 XX WPI; 1999-254290/21.
 DR
 XX P-PSDB; AAY06965.
 XX
 XX Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX
 XX Disclosure; Fig 25A; 55pp; English.
 PS
 XX The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are

CC used to detect E. chaffeensis in patients and E. canis in dogs.

XX Sequence 882 BP; 306 A; 137 C; 140 G; 299 T; 0 other;

Alignment Scores:
 Pred. No.: 5.8e-41 Length: 882
 Score: 448.00 Matches: 100
 Percent Similarity: 55.99% Conservative: 59
 Best Local Similarity: 35.21% Mismatches: 103
 Query Match: 29.67% Indels: 22
 DB: 20 Gaps: 7

US-09-846-808-1 (1-284) x AAX34765 (1-882)

Qy 13 ValMetLeuLeuPheThrProHisLeuAlaSerValLeuAsnHisAsn 32
 Db 40 TTAGTATGCTTATGTCATTAACCTATATATCTCCCAAGGCCATAAACAATACCT 99
 Qy 33 SerMetTyrValGlyIle-----GlnTyrLysProAlaArgGlnHisLeuSer 48
 Db 100 AAAAGTACTACGGATATATATCAGTGGGACAAATATAAACCCAGTGTCTCTGTTTCAGT 159
 Qy 49 LysLeuLeuLeuLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysAsp 68
 Db 160 AATTTCAGTTAAAGAAACCAATGTCTATACTAAACCTTATAGCTTTTAAAAAAGAT 219
 Qy 69 Leu-----LeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPhe 83
 Db 220 GTTGACTCTATTGAACCAACAGCTGATGCCAGTGTAGTATTAGTAACCCATCAATTTT 279
 Qy 84 AsnIleLysTyrAsnProTyrTyrGluAsnArgLeuGlyPheSerGlyIlePheGly 103
 Db 280 ACTATCCCTATACAGCTGTATTTCAAGATAATCTGTCAATTTCAATGGAATATTGGT 339
 Qy 104 TyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIle 123
 Db 340 TACACCTTGTCTGAAGGTACAGAGTTGAAATAGAAGTCTCTTATGAGGAATTTGATGTT 399
 Qy 124 LysAsn---AsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGlu 142
 Db 400 AAAACCTGGAGCTATACACTAAGTGATGCCCTATCGCTATTTTGGATTAGCAGTGAA 459
 Qy 143 IleSerGlyGlySerAsnAsnPro-----AlaAsnAsnLysTyrValThrLeu 158
 Db 460 ATGAAGGTATATAGTTTACACCTTAAAGAAAGTTCTTAATAGTATTTTTCACACTGTA 519
 Qy 159 Ile---AsnAsnGlyIleSerLeuThrSerAlaLeuLeuAsnValCysTyrAspValAsp 177
 Db 520 ATGAGAATGATGATATCTATTAATATCTCTTATAGTAAATGTTGCTACGATTTCTCT 579
 Qy 178 GlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAsp 197
 Db 580 TTGAACAATTTGTCAATATCCCTTACATATGTGGAGGAGCAGGGGTAGATGCTATAGAA 639
 Qy 198 PheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThr 217
 Db 640 TTCCTCGATGATPACACATTAAGTTTGATATCAATCAAGCAAGCTAGGTATTGCTATTCT 699
 Qy 218 ValSerProGlnValSerValPheIleGluTyrTyrHisGlyLeuPheGlyLysLys 237
 Db 700 CTACCATCAACATTAGTCTCTTGGTAGTTTATATACCAATAAGTAATGGCAATCAA 759
 Qy 238 PheGluLysIleProValAsnTyrProCysAspTyrProSerProThrProAsnSer 257
 Db 760 TTTAAATTTAATGTCACACATGTCGTAACCTTGA-----AGT 801
 Qy 258 LysProHisValHisThrThrAlaLeuMetLeuSerIleGlyTyrTyrGlyGlySer 277
 Db 802 ATACCTAAAATT---ACATCGCGAGTTGCTACACTTAATATTGTTATTTTGGAGGTGAA 858
 Qy 278 IleGlyLysLys 281
 Db 859 ATGGTGCAGA 870

RESULT 7
 AAX34749
 ID AAX34749 standard; DNA; 894 BP.

XX AAX34749;

XX 05-JUL-1999 (first entry)

XX DNA encoding OMP-1A protein.

XX Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30; detection; dog; ss.

XX Ehrlichia chaffeensis.

XX WO9913720-A1.

XX 25-MAR-1999.

XX 18-SEP-1998; 98WO-US19600.

XX 19-SEP-1997; 97US-0059353.

XX (OHIS) UNIV OHIO STATE.

XX Ohashi N, Rikihisa Y;

XX WPI; 1999-254290/21.

XX P-PSDB; AAY06949.

XX Novel outer membrane proteins from Ehrlichia chaffeensis and Ehrlichia canis

XX Disclosure; Fig 9A; 55pp; English.

XX The invention provides isolated outer membrane proteins (OMP) from Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part of the OMP family and consist of proteins OMP-1, -1(B to Z) shown in AAY06943-958. The E. canis proteins form part of the p30 family and consist of proteins shown in AAY06959-970. The proteins and genes are used to detect E. chaffeensis in patients and E. canis in dogs.

XX Sequence 894 BP; 308 A; 143 C; 152 G; 291 T; 0 other;

Alignment Scores:

Pred. No.: 9.95e-41 Length: 894
 Score: 446.00 Matches: 102
 Percent Similarity: 55.15% Conservative: 64
 Best Local Similarity: 33.89% Mismatches: 107
 Query Match: 29.54% Indels: 28
 DB: 20 Gaps: 11

US-09-846-808-1 (1-284) x AAX34749 (1-894)

Qy 1 MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuIle-----LeuPhe 18
 Db 10 CTCATGATATAAGAAACAATCTCTTACATAAGTACAGCAATGGTATGCTTATTGTTA 69
 Qy 19 ThrProHisIleSerLeuAlaSerValLeuAsnAsp-----HisAsnSerMet 34
 Db 70 TTACTGTGTATCATCTTTTCAGAAACTATAAACAACAGTGTCTAAAAACAGCCTGGTTA 129
 Qy 35 TyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuLeuLysGlu 54
 Db 130 TATATCAGTGGCGCATACAAACCTAGTGTTCAGTTTTTAGTAATTTTTCATAAAGAA 189
 Qy 55 SerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeu-----70
 Db 190 ACTAATGTTCACACACAGCAGTTAATAGCACTTAAAAAGACATTAATCTCTTCAGTT 249
 Qy 71 ---AsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIleLysTyrAsnPro 89

```
Db 250 GGTAGTAATGCTACACAGGTATTAGCAATCCAGGTAATTCACAAATTCCTTATACGTCA 309
Qy 90 TyrtYrGluAsnAsnArgLeuGlyPheSerGlyIlePheGlyTyrtYrAsnLysAsn 109
Db 310 GAATTTCAAGATAATCTCCCAATTCATCGGCTGTGGTTACTCTTTCTCTGATAGT 369
Qy 110 PheArgIleGluSerGluLeuSerTyrtYrGluThrPheHisIleLysAsn---AsnGlyTyrt 128
Db 370 CTAAGAATTTGAATAGAGGATTTTCATGAAAAATTTGATGTCAAAAACCCCTGAGGTATC 429
Qy 129 LysArgIle---AspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlySer 147
Db 430 ACACAGTAAGTAAGCGGTACCGTATTATTTGACATGACGAGTAAATTAAGATGCGTTC 489
Qy 148 AsnAsnProAla-----AsnAsnLysTyrtYrValThrLeuIle---AsnAsnGlyIle 163
Db 490 TTTGAACCTTAAGCGGAAGATACAGGTGTTTATCATCTACTGTTATGAAAAATGATGATTA 549
Qy 164 SerLeuThrSerAlaLeuIleAsnValCysTyrtYrAsp-----ValAspGlyLeuLysHis 181
Db 550 TCTATTTTATCTACTATGTTAAACGCTCTGTACGATTTTCTGTAGATGAATTACCA--- 606
Qy 182 AsnIleIleThrTyrtYrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLys 201
Db 607 ---GTCCTTACCTTATATATGTCGAGGTATGGGTATAAAGCCCATAGAAATTCCTCGACGCT 663
Qy 202 TyrtYrThrLysPheSerTyrtYrGlnGlyLysLeuGlyAlaSerTyrtYrValSerProGln 221
Db 664 TTACATGTAAATTTGCTTACCAAGCAACATGAGTATTAGTATCACTATTACTATAA 723
Qy 222 ValSerValPheIleGluGlyTyrtYrHisGlyLeuPheGlyLysPheGluLysIle 241
Db 724 GTAAATTTTATCTGATGGGTATTACCATCAAGTAATAGGCAATCAATTCAAAACTTA 783
Qy 242 ProValAsnTyrtYrProCysAspTyrtYrProSerProThrProProAsnSerLysProHisVal 261
Db 784 AAGTAAACCATGTTTACACACTT-----AAAGAAATCTCCTAAAGTC 825
Qy 262 HisThrThrAlaLeuAlaMetLeuSerIleGlyTyrtYrGlySerIleGlyIleLys 281
Db 826 ---ACATCTGCAGTAGCTACACTTGACATGTCATCTTTGGTGGCGAAGTTGGAATAAGA 882
Qy 282 Phe 282
Db 883 TTC 885

RESULT 8
AAx34754
ID AAX34754 standard; DNA; 840 BP.
XX
AC AAX34754;
XX
XX 05-JUL-1999 (first entry)
XX
DE DNA encoding OMP-IV protein.
XX
KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
KW detection; dog; ss.
XX
XX Ehrlichia chaffeensis.
OS
XX W09913720-A1.
XX
XX 25-MAR-1999.
XX
XX 18-SEP-1998; 98WO-US19600.
XX
XX 19-SEP-1997; 97US-0059353.
XX
XX (OHIS ) UNIV OHIO STATE.
XX
XX Ohashi N, Rikihisa Y;
XX
```

DR WPI; 1999-254290/21.
DR P-PSDB; AAY06954.

XX Novel outer membrane proteins from Ehrlichia chaffeensis and
PT Ehrlichia canis
XX

PS Disclosure; Fig 14A; 55pp; English.

XX The invention provides isolated outer membrane proteins (OMP) from
CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
CC in AAY06943-958. The E. canis proteins form part of the P30 family and
CC consist of proteins shown in AAY06959-970. The proteins and genes are
CC used to detect E. chaffeensis in patients and E. canis in dogs.

XX Sequence 840 BP; 279 A; 128 C; 136 G; 297 T; 0 other;

Alignment Scores:

Pred. No.: 3,48e-39 Length: 840
Score: 432.00 Matches: 99
Percent Similarity: 52.22% Conservative: 54
Best Local Similarity: 33.79% Mismatches: 108
Query Match: 28.61% Indels: 32
DB: 20 Gaps: 8

US-09-846-808-1 (1-284) x AAX34754 (1-840)

Qy 5 SerAsnArgLysPheVal-----LeuTrpValMetLeuIleLeuPheThrPro 20
Db 4 AGCAAAAAGAGTTTATTACAATAGGACAGTACTTGCATCTCTATTATCAATCTTATCT 63

Qy 21 HisIleSerLeuAlaSerValLeuAsnAspHis-----AsnSerMetTyrt 35
Db 64 ATTGAATCCTTTTCAGCTATAATCATATCATACAGGAAATAACACTAGTGGTATATAT 123

Qy 36 ValGlyIleGluTyrtYrLysProAlaArgGlnHisLeuSerLysLeuLeuIleGluSer 55
Db 124 ATTACAGGCGCATATAGACAGGAGTATCCCATTTTAGCAATTTCTCAGTAAAGAAACT 183

Qy 56 AlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThr 75
Db 184 AATGTTGATCAATACACTAGTAGGATATAAANAAGT-----GGCTCT 228

Qy 76 GlyIleLysAspAsnThr-----AsnPheAsnIleLysTyrtYrAsnProTyrtYrGluAsn 93
Db 229 TCTATCGATCCTAACACTTATTCAAACTTTCAGGTCCTATATACTCTTACATTTCAGAT 288

Qy 94 AsnArgLeuGlyPheSerGlyIlePheGlyTyrtYrAsnLysAsnPheArgIleGlu 113
Db 289 AATGCTGCTAGTTTTCAGTGGAGCAATTTGGATATTTACCCCGAAAGTCTAGACTTTGAA 348

Qy 114 SerGluLeuSerTyrtYrGluThrPheHisIleLysAsn---AsnGlyTyrtYrLysArgIleAsp 132
Db 349 CTTGAAGGTTCTTACGAAAAATTTGATGTCAAGATCCCTAAAGACTACTCAGCAAAAGAT 408

Qy 133 CysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsn 152
Db 409 GCTTTTAGGTTTTTGTCTAGCAGCAATACGCTCT-----ACTACTGCTTCGTGATGCT 462

Qy 153 AsnLysTyrtYrValThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnVal 172
Db 463 CAAAAATATACAGTTATGAAGAATAATGCTTATCTGTTGCATCAATCATGATCAATGGT 522

Qy 173 CysTyrtYrAspValAspGlyLeuLysHisAsnIleIleThrTyrtYrSerCysLeuGlyPheGly 192
Db 523 TGTATGATCTATCTTTTAAATAATTTAGTCGTATACCTTATATATATGTCAGGATTTGGT 582

Qy 193 ValAspThrIleAspPheLeuSerLysTyrtYrThrLysPheSerTyrtYrGlnGlyLysLeu 212
Db 583 GAAGATTTCATTGAATTTTGTGATCTATCTTTGATCTTTGATCTTTTATCAAGGAAACTA 642

Qy 213 GlyAlaSerTyrtYrValSerProGlnValSerValPheIleGluGlyTyrtYrHisGly 232
Db 213 GlyAlaSerTyrtYrValSerProGlnValSerValPheIleGluGlyTyrtYrHisGly 232

Db 643 GGTATTAGTATTACTTCTTCCCTAAGATTAAATGATTTTGGTGGGTACTATCATAGA 702
 Qy 233 LeuPheGlyLysLysPheGluLysLeuProValAsnTyrProCys-----AspTyr 249
 Db 703 GTTATAGGAATAAATTAAAAATTAAATGTTAACCATGTTGTACACTTGTGATGAATTT 762
 Qy 250 ProSerProThrProProAsnSerLysProHisValHisThrAlaLeuAlaMetLeu 269
 Db 763 CCTAAGCA-----ACTTCTGCAGTAGCTACACTT 792
 Qy 270 SerIleGlyTyrTyrGlyGlySerIleGlyLysPhe 282
 Db 793 AATGTTGCTATTATTTGGTGGTGAAGCTGGAGTAAAGTTT 831
 RESULT 9
 AAX34744
 ID AAX34744 standard; DNA; 852 BP.
 XX AC AAX34744;
 XX 05-JUL-1999 (first entry)
 DT DNA encoding OMP-1B protein.
 DE Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 KW detection; dog; ss.
 XX Ehrlichia chaffeensis.
 OS
 XX WO9913720-Al.
 XX 25-MAR-1999.
 XX 18-SEP-1998; 98WO-US19600.
 XX 19-SEP-1997; 97US-0059353.
 XX (OHIS) UNIV OHIO STATE.
 XX Ohashi N, Rikihisa Y;
 XX WP1; 1998-254290/21.
 DR P-PSDB; AAY06944.
 XX Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX Disclosure; Fig 4A; 55pp; English.
 XX The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX Sequence 852 BP; 301 A; 155 C; 157 G; 239 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 4.05e-39 Length: 852
 Score: 431.50 Matches: 105
 Percent Similarity: 51.90% Conservative: 45
 Best Local Similarity: 36.33% Mismatches: 100
 Query Match: 28.58% Indels: 39
 DB: 20 Gaps: 12
 US-09-846-808-1 (1-284) x AAX34744 (1-852)
 Qy 13 ValMetLeuLeuLeuPheThrProHisLysSerLeuAla----- 25
 Db 37 ATTTCATTAAATGCAATCTTACCTACCAATCTTTTCAGATCTCTGAACCTCAAAATGAT 96
 Qy 26 SerValLeuAsnAsp---HisAsnSerMetTyrValGlyLysGlnTyrLysProAlaArg 44

Db 97 ACAGGAATCAACGACGACGAGAGGCTTCACATTAGTGTAAAGTATAATCCAGCAT 156
 Qy 45 GlnHisLeuSerLysLeuLeuLysGluSerAlaAlaAsn-----Thr 59
 Db 157 TCACACTTTCAGAAATTTCTCAGCTGAAGAAGCTCCCATCAATGGAATCTCTCATCT 216
 Qy 60 ValGluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyLysAsp 79
 Db 217 AAAAAGGTTTTCGGGCTGAAAAAGAC-----GGAGATATAGACAA 258
 Qy 80 AsnThrAsnPheAsnIleLysTyrAsnPro-----TyrTyrGluAsnAsnArgLeu--- 96
 Db 259 TCTCGAATTTTAAAC---AGGACAGATCCAGCCCTCGAGTTTCAGATAAACCTAATATCA 315
 Qy 97 GlyPheSerGlyLysPheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeu 116
 Db 316 GGATTTCTCAGGAAGTATTGGTTATGCTATGGAT---GGGCCAAGAATAAGAACTTGAAGCT 372
 Qy 117 SerTyrGluThrPheHisLys-----AsnAsnGlyTyrLysArgIleAspCys 133
 Db 373 GCATACCAAAATTTGATGCAAAATCTGACAAATGACACATAAGCGGTGACTAC 432
 Qy 134 GluLysHisPheAlaLeuAlaLysGluLysSerGlyGlySerAsnAsnProAlaAsn 153
 Db 433 TATAAATCTTTGGACTATCTCGTCAA-----GACCAATAGCAGATAAG 477
 Qy 154 LysTyrValThrLeuIleAsnAsnGlyLysSerLeuThrSerAlaLeuIleAsnValCys 173
 Db 478 AAATATGTTGCTTAAAAATGAAGCATCACTTTTATGTCATTATGTTAAGCTTGC 537
 Qy 174 TyrAspValAspGlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyVal 193
 Db 538 TATGACATTACAGCTGAAGGAGTACCTTTCATCCGTATGATGTCAGGTGAGGACA 597
 Qy 194 AspThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGly 213
 Db 598 GACCTTATAAAGCTATTTAAGGATTTTAATTTAAATTTCTCATACCAAGGAAATAGGT 657
 Qy 214 AlaSerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeu 233
 Db 658 ATTAGCTATCAATCACACAGAGTTCCTCCGCTTTTATTTGGAGGATACCTACCACGAGTT 717
 Qy 234 PheGlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThr 253
 Db 718 ATAGAAATAATTTTAAACAAATACCTGTA-----ATAACCTGTA 759
 Qy 254 ProProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyr 273
 Db 760 GTATTAGAAGGAGCTCTCTCAAAACC---ACATCTGCGCTAGTAACTATTGACACTGGATAC 816
 Qy 274 TyrGlyGlySerIleGlyLysPhe 282
 Db 817 TTTCGGGAGAGTGGAGTTC 843
 RESULT 10
 AAX34767
 ID AAX34767 standard; DNA; 891 BP.
 XX AC AAX34767;
 XX 05-JUL-1999 (first entry)
 DT DNA encoding P30-7 protein.
 DE Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 KW detection; dog; ss.
 XX Ehrlichia canis.
 OS
 XX WO9913720-Al.
 XX 25-MAR-1999.

XX 18-SEP-1998; 98WO-US19600.
 XX 19-SEP-1997; 97US-0059353.
 XX (OHIS) UNIV OHIO STATE.
 PA Ohashi N, Rikihisa Y;
 PI WPI: 1999-254290/21.
 DR P-PSDB; AAY06967.
 XX Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX Disclosure; Fig 27A; 55pp; English.
 CC The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the p30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX Sequence 891 BP; 327 A; 135 C; 130 G; 299 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 3.16e-37 Length: 891
 Score: 415.00 Matches: 99
 Percent Similarity: 49.68 Conservative: 55
 Best Local Similarity: 31.94 Mismatches: 114
 Query Match: 27.48 Indels: 42
 DB: 20 Gaps: 8
 US-09-846-808-1 (1-284) x AAX34767 (1-891)
 Qy 1 MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuLeuPheThr--- 19
 Db 1 ATGGGAAATCTATGAATAATAAGTCAATCTTAATAAGATTATATTTTAAACATGC 60
 Qy 20 -----ProHisIleSerLeuAlaSerValLeuAsnAsp---HisAsnSerMet 34
 Db 61 ATGCTGTCATTACCTTAATATATCTCTTCAAAAGTAATAACGAAACAACTCTGGTTTG 120
 Qy 35 TyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGlu 54
 Db 121 TATATTAGCGGCATACAAACCCAGTCTTCTGTTTCAGTAATTTTCAGTTAAAGAA 180
 Qy 55 SerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeu----- 69
 Db 181 ACCAACTTTCATACAAACATCTCATAGCTCTTAAACAAGATGTTGATTCGTGAAATT 240
 Qy 70 ---LeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIleLysTyrAsn 88
 Db 241 GATACCTGGTAGTAATACAGCGGTATTAGTAAACCATCTAACTTACAAATCCCTTACT 300
 Qy 89 ProTyrTyrGluAsnAsnArgLeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsnLys 108
 Db 301 GCAGAATTTCAAGACACCACTACTAATGCAATGCTCTATTGGTTATGCTTCTGCTGAA 360
 Qy 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn---AsnGly 127
 Db 361 GGTCCAGAATTTGAATAGAAATTTATCATATGAAATAATTTGATGTTAAATAATCCACAGG 420
 Qy 128 TyrLysArgIle---AspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGly 146
 Db 421 TATACTACAGTAAGAAGTCTTATAGATACCTTGTCTTAGCAGCGTAATAAATTTCT 480
 Qy 147 SerAsnAsnProAlaAsnAsnLys-----TyrValThrLeuIleAsnAsn 161
 Db 481 CTATTCCAAACCAACAAACAAAGAGTAGTGGAATTTACCATGTCGTAAATGAANAACGAT 540
 Qy 162 GlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHis 181

Db 541 GGGTTATCTATCTTATCCAATATAGTAAATATTTGCTACGATTTTCTTAAATAATTTA 600
 Qy 182 AsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLys 201
 Db 601 CCTATATCACCTTATTTATGCGGAGGAATGGGTATAAATGCCATAGAATTTCTTGACGCT 660
 Qy 202 TyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGln 221
 Db 661 TTACATGTGAATTTGCTTATCAAGACGAGGAGGAATAGTTATCAACTATATACGTAA 720
 Qy 222 ValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGlyLysIle 241
 Db 721 ATCACTTATTTATGATGATATATTTACTACGAAGTAATAAGTAATAAATTTAAACCTG 780
 Qy 242 ProValAsnTyrProCysAspTyrProSerProThrProProAsnSerLysProHisVal 261
 Db 781 AAGTCCAA-----CATGTA 795
 Qy 262 His-----ThrThrAlaLeuAlaMetLeuSerIleGlyTyr 273
 Db 796 CATGAACCTTAAAGATAATCCAAAGCTCACATCTGCAGTTGCTACATTCATATAGCATAT 855
 Qy 274 TyrGlyGlySerIleGlyIleLysPheIle 283
 Db 856 TTTGGTAGTGAAGCTGGCATAAGAATTATA 885
 RESULT 11
 AAX34753
 ID AAX34753 standard; DNA; 888 BP.
 XX
 AC AAX34753;
 XX
 DT 05-JUL-1999 (first entry)
 XX
 DE DNA encoding OMP-1U protein.
 XX
 KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; p30;
 KW detection; dog; ss.
 XX
 OS Ehrlichia chaffeensis.
 XX
 PN WO9913720-A1.
 XX
 PD 25-MAR-1999.
 XX
 PF 18-SEP-1998; 98WO-US19600.
 XX
 PR 19-SEP-1997; 97US-0059353.
 XX
 PA (OHIS) UNIV OHIO STATE.
 XX
 PI Ohashi N, Rikihisa Y;
 XX
 DR WPI: 1999-254290/21.
 DR P-PSDB; AAY06953.
 XX
 PT Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX
 PS Disclosure; Fig 13A; 55pp; English.
 CC The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the p30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX
 SQ Sequence 888 BP; 323 A; 127 C; 132 G; 306 T; 0 other;
 Alignment Scores:
 Pred. No.: 3.59e-37 Length: 888

Score: 414.50 Matches: 96
Percent Similarity: 49.50% Conservative: 54
Best Local Similarity: 31.68% Mismatches: 124
Query Match: 27.45% Indels: 29
DB: 20 Gaps: 6

US-09-846-808-1 (1-284) x AAX34753 (1-888)

```
QY 1 MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuLeuLeuPheThrPro 20
    |||:||||| ||| |||: |||: |||: |||
Db 1 ATGACAAGAAATTAATTTGTAATGTTATTAACATTTTGTATTCTTTCCCA 60
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 21 HisIleSerLeuAlaSerValLeuAsnAspHisAsn-----SerMetTyr 35
    ||| ||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 61 CTTAAGTCATTTACACATATGCAATAATAACACATCACTCAAAAAGTTGGATTGTC 120
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 36 ValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSer 55
    :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 121 ATAAGTGGTCAATATAGCCCAAGTATTCCTCATTTCAAGAATTTTTCAGTAGAAGAAAT 180
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 56 AlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThr 75
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 181 ---GACAAAGTAGATGTTAGTCTTACAACTGATGTTACATATATACAGAACAT 237
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 76 GlyIleLysAspAsnThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArg 95
    :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 238 ATATTACGAGATAATACAAAATTCACACTCATATATTGCAAGTTCAAGAACATTTT 297
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 96 LeuGlyPheSerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGlu 115
    :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 298 ATAAATTTACAGTAGTGCAATGTTGTTATTCTGGCAAGGACCAAGGTTAGAAATAGAA 357
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 116 LeuSerTyrGluThrPheHisIleLysAsn---AsnGlyTyrLysArgIleAspCysGlu 134
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 358 AGCTCTTATGGGATTTGATGTTGTAAATATATAAAATTTATGCAAGTACAAAGATGTTAAT 417
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 135 LysHisPheAlaLeuAlaLysGluIleSerGlyGlySer----- 147
    :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 418 AGATATTTCGTTAGTACGTGAAAAAATGGTTCAAATTTCTCTCAAAACCATGAA 477
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 148 -----AsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsn 161
    :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 478 ACTAGTCAACCCCTCGACAGTAATCTCTAAAAGTCTTTTATATCTTTAATGAAGATAAT 537
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 162 GlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHis 181
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 538 GGGGTATTTGTCATCAGTAATAAATCAACGGTGTATGATTTTCTTTTATAACACA 597
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 182 AsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLys 201
    ||| ||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 598 ACAATATCACCTTACGTTATGATAGGAGTTGGAGGAGATTTTATAGAGTTTGTGAAGTA 657
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 202 TyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGln 221
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 658 ATGCATATCAAGTTGCTGCCAAGTAAGTTGGTATTAGCTATCAATATCTCCCTCT 717
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 222 ValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLysIle 241
    :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 718 ATTACTATTTTCTGCTGATGCATATATACAAAGTCAATAATAATTAATTAACACCTA 777
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 242 ProValAsnTyrProCysAspTyrPro---SerProThrProProAsnSerLysProHis 260
    ||| ||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 778 CATGTTAAGTATTATCATATGAATTAATAAACTCACCTACC----- 816
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 261 ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrGlyGlySerIleGlyIle 280
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 817 ---ATTACCTCTGCAACAGCCAACTAAACATTTGAATATTTTGGTGGTGAAGTGGGATG 873
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 281 LysPheIle 283
    :||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 874 AGATTTATA 882
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
```

RESULT 12

AAC68705
ID AAC68705 standard; DNA; 843 BP.
XX
AC AAC68705;
XX
DT 02-MAR-2001 (first entry)
XX
DE Ehrlichia chaffeensis VSA4 gene partial coding sequence.
XX
KW Ehrlichia chaffeensis; VSA4; variable surface antigen 4; MAP1;
KW major antigenic protein 1; antirickettsial; vaccine; gene therapy;
KW Rickettsia; Cowdria; Anaplasma; map2; lhworf3; 4hworf1; 18hworf1;
KW 3gdorf3; ds.
XX
OS Ehrlichia chaffeensis.
XX
PN WO200065063-A2.
XX
PD 02-NOV-2000.
XX
PF 21-APR-2000; 2000WO-US10886.
XX
PR 22-APR-1999; 99US-0130725.
XX
(UYFL) UNIV FLORIDA.
XX
PA Barbet AF, Bowie MV, Ganta RR, Burridge MJ, Mahan SM, McGuire TC;
PI Rurangirwa FR, Moreland AL, Simbi BH, Whitmire WW, Alleman AR;
PI
XX
DR WPI; 2000-679675/66.
DR P-PSDB; AAB36188.
XX
PT New polynucleotides useful as DNA vaccines for conferring immunity to
PT rickettsial infection e.g. heartwater caused by Cowdria ruminantium,
PT comprises major antigenic protein 1 or 2 gene of rickettsial pathogens
PT
XX
PS Claim 4; Page 41; 63pp; English.
XX
CC The present sequence shows a high degree of similarity to the major
CC antigenic protein 1 (MAP1) gene of Ehrlichia sp. The MAP1 gene may be
CC used in a vaccine to protect animals or humans against rickettsial
CC diseases caused by a organisms of Rickettsia sp., Ehrlichia sp.,
CC Anaplasma sp. or Cowdria sp. The vaccine elicits an immune response
CC protective against the rickettsial pathogen. The nucleic acid vaccines
CC can be driven by the human cytomegalovirus (HCMV) enhancer-promoter
CC Cowdria ruminantium genes designated map 2, lhworf3, 4hworf1, 18hworf1
CC and 3gdorf3 may be used in therapeutic and diagnostic applications. The
CC polypeptides are useful for detecting antibodies associated with
CC infection by a rickettsial pathogen whilst the polynucleotides may be
CC used to detect the presence of rickettsial nucleic acids.
XX
SQ Sequence 843 BP; 282 A; 127 C; 159 G; 275 T; 0 other;

Alignment Scores:
Pred. No.: 7,26e-37 Length: 843
Score: 411.50 Matches: 103
Percent Similarity: 52.08% Conservative: 47
Best Local Similarity: 35.76% Mismatches: 113
Query Match: 27.25% Indels: 25
DB: 21 Gaps: 11

US-09-846-808-1 (1-284) x AAC68705 (1-843)

QY 7 ArgLysPheValLeu-----TrpValMetLeuLeuLeuPheThrProHisIleSer 23
:||||||| :||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 12 AAAAAATTTTATATAACAACACTACATAGTATCGCTAATATGCTTCTTACCTGGGAATATCA 71
:||||||| :||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
:||||||| :||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
Db 72 TTTTCTGATCGATACAGACAGCAATGTTGGTGGTAAATTTCTATATATCAGTGGGAATAT 131
:||||||| :||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 41 LysProAlaArgGlnHisLeuSerLysLeuLeuLysGluSerAlaAlaAsnThrVal 60

```
Db 132 GTACCAAGTGTTCACATTTTGGCGTATTCTCTGCTAAACAGGAAAGAAAATACACAATC 191
Qy 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
Db 192 GGAGTATTGGATTAAAGCAGATGGATGGCAGCACAATATCTAAATAATCTCCAGAA 251
Qy 81 ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
Db 252 AATACATTAAAGTTCACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 311
Qy 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 312 GCAGGAGCTGTGGTTATTAATGAAT---GGTCCAGAAATAGAGTTAGAAATGTCCTAT 368
Qy 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
Db 369 GAAACATTGGATGGAACACAGGTAATAACTATATAGAAC---GATGCTCAAAATAT 425
Qy 137 PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnLysTyrVal 156
Db 426 TATGCTTTAAACCCATAAC---AGTGGGGAAAGCTAAGCAATGCAGGTGATAAGTTGTT 482
Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
Db 483 TTTCTAAAAAATGAAGACTACTTGTATATATCACTTATGTTGAATGCATGCTATGATGTA 542
Qy 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAsp 194
Db 543 ATAAGTGAAGGAATACCTTCTCT-----CCTTACATATGTGCAGGTGTGGTACTGAT 596
Qy 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
Db 597 TTAATATCCAGTTTGAAGCTATAAACCCTAAAATTTCTTATCAAGGAAAGTTAGGTTG 656
Qy 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
Db 657 AGTTACTCCATAAGCCAGAGCTCTGTTTGTGTCGACATTTTTCATAAGGTGATA 716
Qy 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro 254
Db 717 GGGAAATGAATTCACA-----GATATCTCTGCTATGATACCC 752
Qy 255 ProAsnSerLysProHisValHisThrThraLeuAlaMetLeuSerIleGlyTyrTyr 274
Db 753 AGTACCTCAACTCTCAGAGGTAATCACTTTACTATATAGTAACACTAAGTGTATGCCACTTT 812
Qy 275 GlyGlySerIleGlyIleLysPhe 282
Db 813 GGAGTGGAACTTGGAGGAAGGTTT 836

RESULT 13
AAV07179
ID AAV07179 standard; DNA; 4683 BP.
AC XX
AC AAV07179;
DT XX
DT 14-SEP-1998 (first entry)
DE Ehrlichia chaffeensis VSA genomic locus.
XX
XX MAP1 homologue; variable surface antigen; VSA1; VSA2; VSA3; VSA4;
KW VSA5; rickettsia; DNA vaccine; ss.
XX
XX Ehrlichia chaffeensis.
OS
FH key Location/Qualifiers
FT CDS 3..134
FT /tag= a
FT /note= "VSA1 partial gene"
FT terminator 203..212
FT /tag= b
FT terminator 226..239
```

```
FT -35_signal /*tag= c
FT 349..354 /*tag= d
FT -10_signal 375..380 /*tag= e
FT RBS 430..434 /*tag= f
FT CDS 439..4299 /*tag= g
FT /note= "VSA2"
FT terminator 1300..1309 /*tag= h
FT terminator 1324..1333 /*tag= i
FT misc_feature 1347..1361 /*tag= j
FT /note= "G-rich region"
FT -35_signal 1473..1478 /*tag= k
FT -10_signal 1499..1504 /*tag= l
FT RBS 1554..1558 /*tag= m
FT CDS 1563..2399 /*tag= n
FT /note= "VSA3"
FT terminator 2444..2456 /*tag= o
FT terminator 2455..2477 /*tag= p
FT misc_feature 2491..2505 /*tag= q
FT -35_signal 2618..2623 /*tag= r
FT -10_signal 2644..2649 /*tag= s
FT RBS 2699..2703 /*tag= t
FT CDS 2708..3550 /*tag= u
FT /note= "VSA4"
FT terminator 3586..3596 /*tag= v
FT terminator 3610..3622 /*tag= w
FT misc_feature 3586..3596 /*tag= x
FT /note= "G-rich region"
FT -35_signal 3767..3772 /*tag= y
FT -10_signal 3788..3793 /*tag= z
FT RBS 3847..3851 /*tag= aa
FT CDS 3856..4683 /*tag= ab
FT /note= "truncated VSA5 gene"
XX
XX WO9816554-A1.
XX
XX 23-APR-1998.
XX
XX 17-OCT-1997; 97WO-US19044.
XX
XX 17-OCT-1996; 96US-0733230.
XX (UYFL ) UNIV FLORIDA.
XX
XX Barbet AF, Burrig MJ, Ganta RR, Mahan SM, McGuire TC;
XX Nyika A, Rurangirwa FR;
XX WPI; 1998-251232/22.
XX
XX DR
```

DR P-PSDB; AAW51091-05.
 XX Composition containing nucleic acid encoding rickettsial antigen -
 PT useful for, e.g. stimulating protective immune response in humans or
 XX animals
 XX
 PS Claim 4; Fig 2A-B; 39pp; English.
 XX
 CC This is the DNA sequence of a 4.6 kb genomic locus of Ehrlichia
 CC chaffeensis that was obtained using a PCR cloning strategy based on
 CC identifying genes homologous to the major antigenic protein MAP1
 CC (see AAW51088) of Cowdria ruminantium. It includes 5 very similar
 CC but non-identical open reading frames (ORFs), of which ORF1 is a
 CC partial gene and ORF5 is nearly complete but lacks 5-7 amino acid
 CC codons and a termination codon (see AAW51091-95). Due to their
 CC similarity to MAP1 surface antigen genes of C. ruminantium, the
 CC E. chaffeensis ORFs are designated variable surface antigen (VSA)
 CC genes 1-5. A claimed composition comprises a nucleic acid (see
 CC AAV07176-82) encoding a polypeptide (see AAW51088-99) that elicits a
 CC protective immune response against a rickettsial pathogen. The
 CC nucleic acid is used, in human or veterinary medicine, in vaccines
 CC to protect against Rickettsia, Ehrlichia, Anaplasma and Cowdria
 CC species. The nucleic acids are also useful as probes to identify
 CC related sequences, e.g. for identification of organisms and for
 CC diagnosing infection. Use of nucleic acid vaccines avoids the
 CC problem of protein purification associated with protein-based
 CC vaccines. The nucleic acid does not replicate in the host but
 CC remains episomal and capable of expressing polypeptide for at least
 CC 19 mth.
 XX
 SQ Sequence 4683 BP; 1590 A; 753 C; 818 G; 1522 T; 0 other;

Alignment Scores:
 Pred. No.: 8,61e-36 Length: 4683
 Score: 411.50 Matches: 103
 Percent Similarity: 52.08% Conservative: 47
 Best Local Similarity: 35.76% Mismatches: 113
 Query Match: 27.25% Indels: 25
 DB: 19 Gaps: 11

US-09-846-808-1 (1-284) x AAV07179 (1-4683)

QY 7 ArgLysPheValLeu-----TrpValMetLeuLeuPheThrProHisIleSer 23
 DB 2717 AAAAAATTTTATATAACAACACTACATAGTAGTCGTAATGCTCTTCTACCTGGAATATCA 2776
 QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
 DB 2777 TTTTCTGATGCGAGTACAGAACGACAAATGTTGGTGGTAATTTCTATATATCAGTGGGAAATAT 2836
 QY 41 LysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSerAlaAlaAsnThrVal 60
 DB 2837 GTACCAAGTGTTCACATTTTGGCGTATTCTCTGCTAACAGGAAAGAAATAACAACATC 2896
 QY 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
 DB 2897 GGAGTATTGGATTAAAGCAAGATGGGATGGCAGCACAAATATCTAAATAATCTCCAGAA 2956
 QY 81 ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
 DB 2957 AATACATTTTAAACGTTCCCAATTTATTCATTAAATATGAAATAATATCCATTTCTAGTGT 3016
 QY 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuLeuSerTyr 118
 DB 3017 GCAGGAGCTGTGGTTATTATGAAT---GGTCCAAGTAATAGATTAGAAATGCTCTAT 3073
 QY 119 GluThrPheHisLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
 DB 3074 GAAACATTTGATGTGAAAAACAGGCTAATAACTATAAGAAC---GATGCTCAACAATAT 3130
 QY 137 PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
 DB 3131 TATGCTTTAACCCATAAC---AGTGGGGGAAAGCTAAGTAATGCAAGTGTATAAGTTTGT 3187

QY 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
 DB 3188 TTTCTAAAAAATGAAGACTACTTGTATATATCATCTTATGTTGAATGCATGCTATGATGA 3247
 QY 177 -----AspGlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAsp 194
 DB 3248 ATAAGTGAAGGAATACCTTTCTCT-----CCTTACATATGTCGAGGTGTGGTACTGAT 3301
 QY 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
 DB 3302 TTAATATCATCTGTTTGAAGCTATAAACCCTAAATTTCTTATCAAGGAAGATTAGTTTG 3361
 QY 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
 DB 3362 AGTACTCATATAAGCCAGAAAGCTCTGTTTTTGGTGGACATTTTTCATAAGGTGATA 3421
 QY 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro 254
 DB 3422 GGAATGAATTCAGA-----GATATTCTGCTATGATACCC 3457
 QY 255 ProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyr 274
 DB 3458 AGTACCTCAACTCTCACAGGTAATCACTTTACTATATAGTAACACTAAGTGTATGCCACTTT 3517
 QY 275 GlyGlySerIleGlyIleLysPhe 282
 DB 3518 GGAGTGAACCTTGGAGGAAGGTTT 3541

RESULT 14
 AAC68716
 ID AAC68716 standard; DNA; 4683 BP.
 XX
 AC AAC68716;
 XX
 DT 02-MAR-2001 (first entry)
 XX
 DE Ehrlichia chaffeensis 28 kDa gene locus.
 XX
 KW Ehrlichia chaffeensis; MAP1; major antigenic protein 1; antirickettsial;
 KW vaccine; gene therapy; Rickettsia; Cowdria; Anaplasma; map2; ihworf3;
 KW 4hworf1; 18hworf1; 3gdorf3; ds.
 XX
 EH Ehrlichia chaffeensis.
 XX
 PN WO200065063-A2.
 XX
 PD 02-NOV-2000.
 XX
 PF 21-APR-2000; 2000WO-US10886.
 XX
 PR 22-APR-1999; 99US-0130725.
 XX
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Barbet AF, Bowie MV, Ganta RR, Burrridge MJ, Mahan SM, McGuire TC;
 PI Rurangirwa FR, Moreland AL, Simbi BH, Whitmire WW, Alleman AR;
 XX
 DR WPI; 2000-679675/66.
 XX
 PT New polynucleotides useful as DNA vaccines for conferring immunity to
 PT rickettsial infection e.g. heartwater caused by Cowdria ruminantium,
 PT comprises major antigenic protein 1 or 2 gene of rickettsial pathogens
 XX
 PS Example 2; Fig 2A-2B; 63pp; English.
 CC
 CC The present sequence is given in a specification relating to nucleic
 CC acid vaccines which may be used to protect animals or humans against
 CC rickettsial diseases caused by a organisms of Rickettsia sp.,
 CC Ehrlichia sp., Anaplasma sp. or Cowdria sp. The vaccine elicits an
 CC immune response protective against the rickettsial pathogen. The
 CC vaccines comprises the major antigenic protein 1 (MAP1) or major

CC antigenic protein 2 (MAP2) of rickettsial pathogens. The nucleic acid
 CC vaccines can be driven by the human cytomegalovirus (HCMV)
 CC enhancer-promoter. Cowdria ruminatum genes designated map 2, ihworf3,
 CC 4hwof1, 18hwof1 and 3dofr3 may be used in therapeutic and diagnostic
 CC applications. The polypeptides are useful for detecting antibodies
 CC associated with infection by a rickettsial pathogen whilst the
 CC polynucleotides may be used to detect the presence of rickettsial
 CC nucleic acids.

XX
 SQ Sequence 4683 BP; 1576 A; 747 C; 831 G; 1526 T; 3 other;

Alignment Scores:

Pred. No.: 8,61e-36 Length: 4683
 Score: 411.50 Matches: 103
 Percent Similarity: 52.08% Conservative: 47
 Best Local Similarity: 35.76% Mismatches: 113
 Query Match: 27.25% Indels: 25
 DB: 21 Gaps: 11

US-09-846-808-1 (1-284) x AAC68716 (1-4683)

QY 7 ArgLysPheValLeu-----TrpValMetLeuLeuPheThrProHisIleSer 23
 Db 2717 AAAAAATTTTATACAACTACATTAGTATCGCTAATGTCCTTCTTACTCGGAATATCA 2776
 QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
 Db 2777 TTTTCTGATGCAGTACAGACGACATGTTGGTGGTAATTTCTATATCAGTGGGAATAT 2836
 QY 41 LysProAlaArgGlnHisIleSerLysLeuLeuIleLysGluSerAlaAlaAsnThrVal 60
 Db 2837 GTACCAAGTCTTTCACATTTTGGCGTATTCTCTGCTAAACAGCAAGAAATAACAATC 2896
 QY 61 GluValPheGlyLeuLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
 Db 2897 GGAGTATTGGATTAAGCAAGATGGGATGGCAGCACAAATATCTAAAAATTCCTCAGAA 2956
 QY 81 ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
 Db 2957 AATACATTTTAACGTTCCAAATATTATTCATTAAATAGAAATAAATCCATTTCTAGGTTT 3016
 QY 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
 Db 3017 GCAGGAGCTGTTGGTATTATTAATGAAT---GGTCCAGAAATAGAGTAGAATGTCCTAT 3073
 QY 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
 Db 3074 GAACATTTTGATGTGAAACACGAGGTAAATAACTATAAGAAC---GATGCTCACAAATAT 3130
 QY 137 PheAlaLeuAlaLysGluIleSerGlyLysSerAsnAsnProAlaAsnLysTyrVal 156
 Db 3131 TATGCTTTAACCCATAAC---AGTGGGGGAAGCTAAGCAATGCAGGTAGTAAAGTTGTT 3187
 QY 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCystTyrAspVal 176
 Db 3188 TTTCTAAAAAATGAAGGACTACTTGATATATCACTTATCTGAATGCATGCTATGATGA 3247
 QY 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValasp 194
 Db 3248 ATAAGTGAAGGAATACCTTCTCT-----CCTTACATATGTCGAGGTGTTGGTACTGAT 3301
 QY 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
 Db 3302 TTAATATCATGCTTTGAAGCTATAAACCCCTAAATTTCTATCAAGAAAGTTAGGTTTG 3361
 QY 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
 Db 3362 AGTTACTCCATAAGCCAGCAAGCTTCTGTTTGTGGTGGCAATTTTCATAAGGTGATA 3421
 QY 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro 254
 Db 3422 GGGAAATGAATTCAGA-----GATATTCCTGCTATGATACCC 3457

QY 255 ProAsnSerLysProHisValHisThrAlaLeuAlaMetLeuSerIleGlyTyrTyr 274
 Db 3458 AGTACCTCAACTCTCACAGGTATATCATTACTATAGTAACTAACACTAAGTGTATGCCACTTT 3517
 QY 275 GlyGlySerIleGlyIleLysPhe 282
 Db 3518 GGAGTGGAACTTGGAGGAAGGTTT 3541
 RESULT 15
 AAS07578
 ID AAS07578 standard; DNA; 4683 BP.
 XX AAS07578;
 AC AAS07578;
 XX
 DT 23-OCT-2001 (first entry)
 DE DNA encoding variable surface antigens 1-5 (VSA1-5) from E. chaffeensis.
 XX
 KW Major antigenic protein 1; MAP1; vaccine; immunogenic; rickettsial;
 KW infection; heartwater; diagnostic; variable surface antigen; VSA; ds.
 XX
 OS Ehrlichia chaffeensis.
 FH Key Location/Qualifiers
 FT CDS 3..134
 FT /*tag= a
 FT /product= "Variable surface antigen 1 (VSA1)"
 FT /partial
 FT /note= "No start codon"
 FT terminator 179..188
 FT /*tag= b
 FT /note= "Transcription terminator of VSA1 gene"
 FT terminator 202..212
 FT /*tag= c
 FT /note= "Transcription terminator of VSA1 gene"
 FT GC_signal 226..239
 FT /*tag= d
 FT -35_signal 349..354
 FT /*tag= e
 FT /note= "-35 region of VSA2 gene"
 FT -10_signal 375..380
 FT /*tag= f
 FT /note= "-10 region of VSA2 gene"
 FT RBS 430..434
 FT /*tag= g
 FT /note= "Ribosome binding site of VSA2 gene"
 FT CDS 439..1299
 FT /*tag= h
 FT /product= "Variable surface antigen 2 (VSA2)"
 FT terminator 1300..1309
 FT /*tag= i
 FT /note= "Transcription termination signal of VSA2"
 FT terminator 1324..1333
 FT /*tag= j
 FT /note= "Transcription termination signal of VSA2"
 FT GC_signal 1349..1361
 FT /*tag= k
 FT /note= "GC-rich sequence of VSA3 gene"
 FT -35_signal 1473..1478
 FT /*tag= l
 FT -10_signal 1499..1504
 FT /note= "-35 region of VSA3 gene"
 FT /*tag= m
 FT /note= "-10 region of VSA3 gene"
 FT RBS 1554..1558
 FT /*tag= n
 FT /note= "Ribosome binding site of VSA3 gene"
 FT CDS 1563..2399
 FT /*tag= o
 FT /product= "Variable surface antigen 3 (VSA3)"
 FT terminator 2443..2455
 FT /*tag= p

FT /note= "Transcription termination signal of VSA3"
FT 2465..2477
FT /*tag= q
FT /note= "Transcription termination signal of VSA3"
FT 2491..2505
FT /*tag= i
FT /note= "GC-rich sequence of VSA4 gene"
FT 2618..2623
FT /*tag= s
FT /note= "-35 region of VSA4 gene"
FT 2644..2649
FT /*tag= t
FT /note= "-10 region of VSA4 gene"
FT 2699..2703
FT /*tag= u
FT /note= "Ribosome binding site of VSA4 gene"
FT 2708..3550
FT /*tag= v
FT /product= "Variable surface antigen 4 (VSA4)"
FT 3586..3596
FT /*tag= w
FT /note= "Transcription termination signal of VSA4"
FT 3610..3622
FT /*tag= x
FT /note= "Transcription termination signal of VSA4"
FT 2491..2505
FT /*tag= y
FT /note= "GC-rich sequence of VSA5 gene"
FT 3767..3772
FT /*tag= z
FT /note= "-35 region of VSA5 gene"
FT 3788..3793
FT /*tag= aa
FT /note= "-10 region of VSA5 gene"
FT 3847..3851
FT /*tag= ab
FT /note= "Ribosome binding site of VSA5 gene"
FT 3856..4683
FT /*tag= ac
FT /product= "Variable surface antigen 5 (VSA5)"
FT /partial
FT /note= "No stop codon"
XX
PN USG251872-B1.
XX
XX
XX 26-JUN-2001.
XX
XX 17-OCT-1997; 97US-0953326.
XX
XX 17-OCT-1996; 96US-0733230.
XX
XX (UYFL) UNIV FLORIDA.
XX
XX Barbet AF, Ganta RR, McGuire TC, Burrige MJ, Nyika A;
XX Rurangirwa FR, Mahan SM, Bowie MV, Alleman AR;
XX
XX WPI; 2001-424487/45.
XX
XX P-PSDB; AAU04195, AAU04196, AAU04197, AAU04198, AAU04199.
XX
XX New MAP2 genes and polypeptides useful as vaccines for conferring
XX immunity to human and animal rickettsial diseases, e.g. heartwater, or
XX as molecular markers in nucleic acid analysis procedures -
XX
XX Example 3; Fig 2A-2B; 30pp; English.
XX
XX The sequence represents the coding sequence of variable surface antigen
XX (VSA) gene locus encoding VSA1-5 proteins of Ehrlichia chaffeensis, which
XX have similarity to major antigen proteins (MAP). The MAP polynucleotides
XX and polypeptides are useful as vaccines for conferring immunity to
XX rickettsia infection, including Cowdria ruminantium causing heartwater.
XX The MAP polynucleotides may be used as molecular markers in nucleic acid
XX analysis procedures, and to produce the MAP polypeptides, which may
XX be used to raise antibodies that are reactive with the polypeptides.

CC The nucleic acids may further be used as probes to identify
CC complementary sequences within other nucleic acid molecules or genomes,
CC where such probes can be applied to identify or distinguish infectious
CC strains of organisms in diagnostic procedures or in rickettsial
CC research where identification of particular organisms or strains is
CC needed.
XX
XX SQ Sequence 4683 BP; 1578 A; 739 C; 833 G; 1533 T; 0 other;

Alignment Scores:
Pred. NO.: 8.61e-36 Length: 4683
Score: 411.50 Matches: 103
Percent Similarity: 52.08% Conservative: 47
Best Local Similarity: 35.76% Mismatches: 113
Query Match: 27.25% Indels: 25
DB: 22 Gaps: 11

US-09-846-808-1 (1-284) x AAS07578 (1-4683)
Qy 7 ArgLysPheValLeu-----TrpValMetLeuLeuPheThrProHisIleSer 23
Db 2717 AAAAAATTTTATACAACTACATTAGTAGCGCTAATGCTCTTTACCTGGAATATCA 2776
Qy 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValcIyleGlnTyr 40
Db 2777 TTTTCTGATGCAGTACAGACGACAATGTTGGTGCTAATTTCTATATCAGTGGGAATAT 2836
Qy 41 LysProAlaArgGlnHisLeuSerLysLeuLeuLeuLysGluSerAlaAlaAsnThrVal 60
Db 2837 GTACCAAGTGTTCACATTTTGGCGTATTCTCTGCTAAACAGGAAGAAATACAAACATC 2896
Qy 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
Db 2897 GGAGTATTGGATTAAACCAAGATTGGATGGCAGCACACAATATCTAAAAATCTCCAGAA 2956
Qy 81 ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
Db 2957 AATACATTAAACGTTCCAAATTAATTAATTAATAATGAAATAATCCATTTCTAGGTTT 3016
Qy 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 3017 GCAGAGCTGTGGTTATTATTATGAAT---GGTCCAGAAGATAGATTAGAAATGTCCTAT 3073
Qy 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
Db 3074 GAAACATTTGATGTGAAAAACAGGGTAAATAACTATAAGAAC---GATGCTCACAAATAT 3130
Qy 137 PheAlaLeuAlaLysGluIleSerGlySerAsnAsnProAlaAsnLysTyrVal 156
Db 3131 TATGCTTTAAACCCATAAC---AGTGGGGGAAAGCTTAAGCAATGCAGGTGATAAGTTGTT 3187
Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCystTyrAspVal 176
Db 3188 TTTCTAAAAAATGAAGGACTACTTGATATATACACTATTGTTGAATGCATGCTATGATGTA 3247
Qy 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAsp 194
Db 3248 ATAAGTGAAGGAATACCTTCTCT-----CCTTACATATGTGCAGGTGTTGGTACTGAT 3301
Qy 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
Db 3302 TTAATATCCATGTTTGAAGCTATAACCCCTAAATTTCTTATCAAGGAAGTAGGTTTG 3361
Qy 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
Db 3362 AGTACTCCATAAGCCCAAGAGCTTCTGTTTGTGTGGACATTTTCATCAAGTGATA 3421
Qy 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro 254
Db 3422 GGAATGAATTCAGA-----GATATTCCTGCTATGATACCC 3457
Qy 255 ProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyr 274

FT /*tag= c
 FT /product= "Mature Eca28SA3 protein (28-kDa)"
 PN
 XX WO200032745-A2.
 XX

PD 08-JUN-2000.
 XX

XX 24-NOV-1999; 99WO-US28075.
 XX

XX 30-NOV-1998; 98US-0201458.
 PR

PR 03-MAR-1999; 99US-0261358.
 XX

XX (RERE-) RES DEV FOUND.
 PA

XX Walker DH, Yu X, McBride JW;
 PI

XX WPI; 2000-412298/35.
 DR

DR P-PSDB; AAY1479.
 DR

XX Ehrlichia canis antigens useful for vaccinating against canine
 PT ehrlichiosis in dogs -

XX Claim 5; Page 67-68; 86pp; English.
 PS

XX The patent relates to homologous 28-kilodalton (kDa) protein genes of
 CC Ehrlichia canis, designated Eca28SA1, Eca28SA2, Eca28SA3, Eca28-1 and
 CC Eca28-2. These genes are members of a polymorphic multiple gene family
 CC and contained in a single locus of 5.592 kb. The 28-kDa proteins are
 CC immunoreactive with anti-E. canis serum hence are important
 CC immunoprotective antigens. The protein is useful for vaccinating
 CC against E. canis infections such as canine ehrlichiosis in dogs.
 CC Canine ehrlichiosis, also known as canine tropical pancytopenia, is a
 CC tick-borne rickettsial disease of dogs. Eca28-1 is conserved amongst
 CC different strains of E. canis and hence useful for serodiagnosis of
 CC canine ehrlichiosis. The present sequence is a DNA encoding E. canis
 CC Eca28SA3 30-kDa protein which is post-translationally modified to a
 CC mature 28-kDa protein by cleavage of N-terminal signal sequence.
 XX

XX SQ Sequence 840 BP; 282 A; 137 C; 150 G; 271 T; 0 other;
 SQ

Alignment Scores:
 Pred. No.: 2,53e-34 Length: 840
 Score: 389.00 Matches: 98

Percent Similarity: 51.09% Conservative: 43
 Best Local Similarity: 35.51% Mismatches: 115

Query Match: 25.76% Indels: 20
 DB: 21 Gaps: 10

US-09-846-808-1 (1-284) x AAD01294 (1-840)

QY 15 LeuileLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn----- 32
 Db 43 TTAATGACTATGCTCCAGCATATCTTTTCTGATCTACTATACAGAGCATACACCTGGT 102

QY 33 SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuLeuIle 52
 Db 103 AGCTTCTACATCAGTGGAAATATATGACCAAGTGTTCACATTTTGGTGTTCACGCT 162

QY 53 LysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeuAsnAsp 72
 Db 163 AAAGAAGAAAGAAACTCAACTGTGGAGTGTTCGATTAACATGATTCGAATGGAGGT 222

QY 73 LeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIleLys---TyrAsnProTyrTyr 91
 Db 223 ACAATATCTAACTTCTCCAGAAATATATTCACATTCGTTCAAAATATTCGTTAAATAC 282

QY 92 GluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrAsnLysAsnAsp 110
 Db 283 GAAACCAACCCATCTTCTAGGTTTGCAGAGCTATTGTTATTCAATGGGT---GGCCCA 339

QY 111 ArgIleGluSerGluLeuSerTyrGluThrPheHisIleLys-----AsnAsnGlyTyr 128
 Db 340 AGAATAGAACTTGAAGTTCTGTACGAGACATTCGATGTGAAATATCAGAAATAATAT 399

QY 129 LysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGlySerAsn 148
 Db 400 AAGAAC---GGCGCACACAGATCTGCTTTATCTCATCATAGTTTCAGCAACAGCATG 456
 QY 149 AsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThrSerAla 168
 Db 457 TCCTCCGCAAGTAACAATTTGTTTCTTAAATAATGAAGGTTAATTTGACTTATCATTT 516
 QY 169 LeuIleAsnValCysTyrAsp-----ValAspGlyLeuLysHisAsnIleIleThrTyr 186
 Db 517 ATGATAAATGCATGCTATGACATAATAATTGAAGGAATGCCTTTTCA-----CCTTAT 570
 QY 187 SerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThrLysPhe 206
 Db 571 ATTTGTGCAGGTGTGGTACTGATGTTGTTCCATGTTTGAAGCATATAAATCCCTAAATTT 630
 QY 207 SerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIle 226
 Db 631 TCTTACCAAGGAAACTAGGATTAGGTTATAGTATAAGTTTCAGAGGCCTCTGTTTTATC 690
 QY 227 GluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGlyLysIleProValAsnTyrPro 246
 Db 691 GGTGGACACTTTCACAGAGTCATAGGTAATGAATTTAGAGACATCCCTGCTATGTTCT 750
 QY 247 CysAspTyrProSerProThrProProAsnSerLysProHisValHisThrThrAlaLeu 266
 Db 751 AGTGA-----TCAATCTTCAGAAACCAA-----TTTGAATA 786
 QY 267 AlaMetLeuSerIleGlyTyrTyrGlySerIleGlyIleLysPhe 282
 Db 787 GTAACACTAAATGTGTCTACCTTTGGCATAGAACTTGGAGGAAGATTT 834

RESULT 18
 AAD01295
 ID AAD01295 standard; DNA; 2037 BP.
 XX
 AC AAD01295;
 XX

DT 12-OCT-2000 (first entry)
 XX

DE Ehrlichia canis immunoreactive protein genes Eca28SA2 and Eca28SA3.
 XX

KW Homologous mature 28-kDa protein gene; Eca28SA2; Eca28SA3; vaccine;
 KW p28 gene; polymorphic multiple gene family; immunoprotective antigen;
 KW antibacterial; canine ehrlichiosis; canine tropical pancytopenia;
 KW tick-borne rickettsial disease; serodiagnosis; immunoreactive; ds.
 XX

OS Ehrlichia canis.
 XX

PH Key Location/Qualifiers
 FT CDS 1..852
 FT /*tag= a
 FT /product= "Eca28SA2 protein (30-kDa)"
 FT misc_feature 853..1194
 FT /*tag= b
 FT /note= "Intergenic non-coding region NC2"
 FT CDS 1195..2037
 FT /*tag= c
 FT /product= "Eca28SA3 protein (30-kDa)"
 FT sig_peptide 1195..1263
 FT /*tag= d
 FT mat_peptide 1264..2034
 FT /*tag= e
 FT /product= "Mature Eca28SA3 protein (28-kDa)"
 XX

XX WO200032745-A2.
 XX

XX 08-JUN-2000.
 PD

XX 24-NOV-1999; 99WO-US28075.
 PF

XX 30-NOV-1998; 98US-0201458.
 PR

PR 03-MAR-1999; 99US-0261358.
 PA (RERE-) RES DEV FOUND.
 PI Walker DH, Yu X, McBride JW;
 DR WPI; 2000-412298/35.
 DR P-PSDB; AAY71478, AAY71479.
 XX Ehrlichia canis antigens useful for vaccinating against canine
 PT ehrlichiosis in dogs -
 XX
 PS Example 15; Fig 7; 86pp; English.
 XX
 CC The patent relates to homologous 28-kiloDalton (kDa) protein genes of
 CC Ehrlichia canis, designated Eca28SA1, Eca28SA2, Eca28SA3, Eca28-1 and
 CC Eca28-2. These genes are members of a polymorphic multiple gene family
 CC and contained in a single locus of 5.592 kb. The 28-kDa proteins are
 CC immunoreactive with anti-E. canis serum hence are important
 CC immunoprotective antigens. The protein is useful for vaccinating
 CC against E. canis infections such as canine ehrlichiosis in dogs.
 CC Canine ehrlichiosis, also known as canine tropical pancytopenia, is a
 CC tick-borne rickettsial disease of dogs. Eca28-1 is conserved amongst
 CC different strains of E. canis and hence useful for serodiagnosis of
 CC canine ehrlichiosis. The present sequence is that of homologous
 CC genes encoding E. canis Eca28SA2 and Eca28SA3 30-kDa proteins which
 CC are post-translationally modified to corresponding
 CC mature 28-kDa proteins by cleavage of N-terminal signal sequence.
 XX
 SO Sequence 2037 BP; 687 A; 321 C; 342 G; 687 T; 0 other;
 Alignment Scores:
 Pred. No.: 9,06e-34 Length: 2037
 Score: 389.00 Matches: 98
 Percent Similarity: 51.09% Conservative: 43
 Best Local Similarity: 35.51% Mismatches: 115
 Query Match: 25.76% Indels: 20
 DB: 21 Gaps: 10
 US-09-846-808-1 (1-284) x AAD01295 (1-2037)
 Qy 15 LeuileLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn----- 32
 Db 1237 TTAATGTACTATGCTCCAGCAATATCTTTCTGATACATATACAGACGAAACACTGGT 1296
 Qy 33 SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuLeuIle 52
 Db 1297 AGCTTCTACATCAGTGGAAATATGACCAAGTGTTCACATTTTGGTGTTCCTCAGCT 1356
 Qy 53 LysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeuAsnAsp 72
 Db 1357 AAAGAAGAAAGAACTCAACTGTTGGAGCTTTTGGATTAAACATGATTGGAATGGAGGT 1416
 Qy 73 LeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIleLys---TyrAsnProTyrTyr 91
 Db 1417 ACAATATCAATCTCTCTCCAGAAATATATTCACAGTTCAAAATATTATTCGTTAAATAC 1476
 Qy 92 GluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsnLysAsnPhe 110
 Db 1477 GAAACAACCCATCTTAGGTTTTCGAGAGCTATTGGTTATTCATGGGT---GGCCCA 1533
 Qy 111 ArgIleGluSerGluLeuSerTyrGluThrPheHisIleLys-----AsnAsnGlyTyr 128
 Db 1534 AGAATAGAAGTGAAGTCTGTACGAGACATTCGATGCGAAATATCAGAACAAATATAT 1593
 Qy 129 LysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyCysSerAsn 148
 Db 1594 AAGAAC---GGCGCACACAGATGCTGCTTATCTCATCATGATTCAGCAACAGCATG 1650
 Qy 149 AsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThrSerAla 168
 Db 1651 TCCTCCGCAAGTAACAAATTTGTTTCTTTAAATAAAGAGGTTTAATTCGACTTATCATTT 1710

Qy 169 LeuileAsnValCysTyrAsp-----ValAspGlyLeuLysHisAsnIleIleThrTyr 186
 Db 1711 ATGATAAATGCATGCTATGACATAATAATTAAGGAATGCCTTTTCA-----CCTTAT 1764
 Qy 187 SerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThrLysPhe 206
 Db 1765 ATTTGTGAGGTGGTGGTACTGATGTTTTCCTCATGTTTGAAGCTATAAATCCTAAAT 1824
 Qy 207 SerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIle 226
 Db 1825 TCTTACCAGGAAACTAGGATTAGTTATAGTATAAGTTACAGACCTCTGTTTTATC 1884
 Qy 227 GluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsnTyrPro 246
 Db 1885 GGTGCACACTTTCACAGAGTCATAGTAAATTTAGAGACATCCTGCTATGGTTCCT 1944
 Qy 247 CysAspTyrProSerProThrProProAsnSerLysProHisValHisThrThrAlaLeu 266
 Db 1945 AGTGA-----TCAAAATCTCCAGAAACCAA-----TTTGCAATA 1980
 Qy 267 AlaMetLeuSerIleGlyTyrTyrGlySerIleGlyIleLysPhe 282
 Db 1981 GTAACACTAAATGTGTCTACCTTGGCATAGAACTGGAGGAAGATTT 2028
 RESULT 19
 ID AAX34747 standard; DNA; 837 BP.
 XX AAX34747;
 AC AAX34747;
 DT 05-JUL-1999 (first entry)
 XX DNA encoding OMP-1E protein.
 DE Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 KW detection; dog; ss.
 XX Ehrlichia chaffeensis.
 OS WO9913720-A1.
 PN 25-MAR-1999.
 PD 18-SEP-1998; 98WO-US19600.
 PF 19-SEP-1997; 97US-0059353.
 PR (OHIS) UNIV OHIO STATE.
 PA Ohashi N, Rikihisa Y;
 PI WPI; 1999-254290/21.
 DR P-PSDB; AAY06947.
 XX Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 PS Disclosure; Fig 7A; 55pp; English.
 CC The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX
 SQ Sequence 837 BP; 280 A; 130 C; 159 G; 268 T; 0 other;
 Alignment Scores:
 Pred. No.: 3.71e-34 Length: 837
 Score: 387.50 Matches: 102
 Percent Similarity: 51.05% Conservative: 44
 Best Local Similarity: 35.66% Mismatches: 117

```
Query Match: 25.66% Indels: 23
DB: Gaps: 10
US-09-846-808-1 (1-284) x AAX34747 (1-837)
QY 7 ArgLysPheValLeu-----TrpValMetLeuLeuPheThrProHisIleSer 23
Db 10 AAAAAATTTTATAACAACGTCATAGTATCATCTTCCTTCTTACCTGGAATATCA 69
QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
Db 70 TTTTCTGATCCAGTCGCAAGTGCATATATTAGTGGTAATTTCTATGTAGTGGCAAGTAT 129
QY 41 LysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSerAlaAlaAsnThrVal 60
Db 130 ATGCCAAGTCCTCGCATTTTGGCATGTTTCTGCCAAGAAGAAAAATCCTACTGTT 189
QY 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuThrGlyIleLysAspAsn 80
Db 190 GCATTGTATGGCTTAAACAAGATTTGGGAAGGATTAGCTCATCAAGTCACAATGATAAT 249
QY 81 ThrAsnPheAsnIleLys---TyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
Db 250 ---CATTTCAATACCAAGGTTATTTCATTTAAATATGAAATAACCCATTTTATAGGTTT 306
QY 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 307 GCAGGAGCTATTGGTATTCAATGGGT---GGTCCAAGACTAGAGTTTGAAGTGCCTAT 363
QY 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
Db 364 GAAACATTTTGACGTTAAAAACAGGTTAATAATGAAATAACATAAAAAAT---GATGCTCAGATAC 420
QY 137 PheAlaLeuAlaLysGluIleSerGlyLysSerAsnAsnProAlaAsnAsnLysTyrVal 156
Db 421 TGTGCTTTAGGTCAACAGACACACAGCGGAATA-----CCTAAAACTAGTAATAATAGTA 474
QY 157 ThrLeuLeuAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
Db 475 CTGTTAAACACGAGGATGCTTGACATATCATTTATGCTAAATGCATCCTATGATATA 534
QY 177 AspGlyLeuLysHisAsnIleLeuThrTyrSerCysLeuGlyPheGlyValAlaSerThrIle 196
Db 535 ATAAACGAGACATACCTTTGCTCCTTACATATGTGCAGGTGTTGGTACTGATTTAATA 594
QY 197 AspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyr 216
Db 595 TCCATGTTTGAAGCTCAAAATCCTAAATTTCTTACCAGGGAAGTTAGTCTTAAGTTAC 654
QY 217 ThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLys 236
Db 655 TCTATAAACCCAGACGCTTCTGTATTATTATGGTGGACATTTTTCATAGGTGATAGGAAC 714
QY 237 LysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrProProAsn 256
Db 715 GAATTTTAGGACATTCCTACTCTGAAGACATTTGTTACGTCATCAGCTACTCCA----- 768
QY 257 SerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGly 276
Db 769 -----GATCTAGCAATAGTAACACTAAGTGTATGCTCATTTTGGGAATA 810
QY 277 SerIleGlyIleLysPhe 282
Db 811 GAACCTGGAGGAAGGTTT 828
RESULT 20
AAC68704
ID AAC68704 standard; DNA; 837 BP.
XX
AC AAC68704;
XX
DT 02-MAR-2001 (first entry)
XX
```

Ehrlichia chaffeensis VSA3 gene partial coding sequence.

XX
KW Ehrlichia chaffeensis; VSA3; variable surface antigen 3; MAP1;
KW major antigenic protein 1; antirickettsial; vaccine; gene therapy;
KW Rickettsia; Cowdria; Anaplasma; map2; lhworf3; 4hworf1; 18hworf1;
KW 3gdorf3; ds.
XX
OS Ehrlichia chaffeensis.
XX
PN W0200065063-A2.
XX
PD 02-NOV-2000.
XX
PF 21-APR-2000; 2000WO-US10886.
XX
PR 22-APR-1999; 99US-0130725.
XX
PA (UYFL) UNIV FLORIDA.
XX
PI Barbet AF, Bowie MV, Ganta RR, Burridge MJ, Mahan SM, McGuire TC;
PI Rurangirwa FR, Moreland AL, Simbi BH, Whitmire WW, Alleman AR;
XX
DR WPI; 2000-679675/66.
DR P-PSDB; AAB36187.
XX
PT New polynucleotides useful as DNA vaccines for conferring immunity to
PT rickettsial infection e.g. heartwater caused by Cowdria ruminantium,
PT comprises major antigenic protein 1 or 2 gene of rickettsial pathogens
PT
XX
PS
XX
CC Claim 4; Page 40; 63pp: English.
CC The present sequence shows a high degree of similarity to the major
CC antigenic protein 1 (MAP1) gene of Ehrlichia sp. The MAP1 gene may be
CC used in a vaccine to protect animals or humans against rickettsial
CC diseases caused by a organisms of Rickettsia sp., Ehrlichia sp.,
CC Anaplasma sp. or Cowdria sp. The vaccine elicits an immune response
CC protective against the rickettsial pathogen. The nucleic acid vaccines
CC can be driven by the human cytomegalovirus (HCMV) enhancer-promoter.
CC Cowdria ruminantium genes designated map 2, lhworf3, 4hworf1, 18hworf1
CC and 3gdorf3 may be used in therapeutic and diagnostic applications. The
CC polypeptides are useful for detecting antibodies associated with
CC infection by a rickettsial pathogen whilst the polynucleotides may be
CC used to detect the presence of rickettsial nucleic acids.
XX
SQ Sequence 837 BP; 279 A; 130 C; 159 G; 269 T; 0 other;

Alignment Scores:
Pred. No.: 3.71e-34 Length: 837
Score: 387.50 Matches: 102
Percent Similarity: 51.05% Conservative: 44
Best Local Similarity: 35.66% Mismatches: 117
Query Match: 25.66% Indels: 23
DB: Gaps: 10

US-09-846-808-1 (1-284) x AAC68704 (1-837)

QY 7 ArgLysPheValLeu-----TrpValMetLeuLeuPheThrProHisIleSer 23
Db 12 AAAAAATTTTATAACAACGTCATAGTATCATCTTCCTTCTTACCTGGAATATCA 71
QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
Db 72 TTTTCTGATCCAGTCGCAAGTGCATATATTAGTGGTAATTTCTATGTAGTGGCAAGTAT 131
QY 41 LysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSerAlaAlaAsnThrVal 60
Db 132 ATGCCAAGTCCTCGCATTTTGGCATGTTTCTGCCAAGAAGAAAAATCCTACTGTT 191
QY 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuThrGlyIleLysAspAsn 80
Db 192 GCATTGTATGGCTTAAACAAGATTTGGGAAGGATTAGCTCATCAAGTCACAATGATAAT 251

Qy 81 ThrAsnPheAsnIleLys---TyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
 Db 252 ---CATTTCAATCAACAAGGGTTATTCAATTAATGAAATACCCATTTTAGGGTTT 308
 Qy 99 SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
 Db 309 GCAGAGAGCTATTGGTTATTCAATGGGT---CGTCCAAGAGTAGAGTTTGAAGTCTCCTAT 365
 Qy 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleLeuAsnGlyLysHis 136
 Db 366 GAAACATTGACGTGTTAAATCAAGGTAATTAATCAATAAAT---GATGTCACAGATAC 422
 Qy 137 PheAlaLeuAlaLysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
 Db 423 TGTGCTTTAGTCAACAGACAGCGGAATA-----CCTAAACACTAGTAAATACGTA 476
 Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
 Db 477 CTGTTAAACCGAAGGATTCCTGCACATATCATTTATGCTAAATGATGCTATGATATA 536
 Qy 177 AspGlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIle 196
 Db 537 ATAAACGAGACATACCTTTGCTCCTTACATATGTCAGGTGTTGGTACTGATTTAATA 596
 Qy 197 AspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyr 216
 Db 597 TCCATGTTTGAAGCTACAAATCCTAAATTTCTTACCAAGGAAAGTTAGGTCTAAGTTAC 656
 Qy 217 ThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLys 236
 Db 657 TCTATAACCCAGAGCTCTGTATTATGTTGGTGACATTTTCATAAGGTGATAGGAAC 716
 Qy 237 LysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrProProAsn 256
 Db 717 GAATTTAGGACATTCCTACTCTGAAAGCATTTGTTAGTCTCATCAGTACTCCA----- 770
 Qy 257 SerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGly 276
 Db 771 -----GATCTAGCAATAGTAACACTAAGTGTATGTCATTTTGGATA 812
 Qy 277 SerIleGlyIleLysPhe 282
 Db 813 CAACTTGGAGGAAGGTTT 830
 RESULT 21
 AAX34762
 ID AAX34762 standard; DNA; 843 BP.
 XX AC AAX34762;
 XX DT 05-JUL-1999 (first entry)
 XX DE DNA encoding P30-2 protein.
 XX KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 XX OS detection; dog; ss.
 XX PN Ehrlichia canis.
 XX PR WO9913720-A1.
 XX PD 25-MAR-1999.
 XX PF 18-SEP-1998; 98WO-US19600.
 XX PR 19-SEP-1997; 97US-0059353.
 XX PA (OHIS) UNIV OHIO STATE.
 XX PI Ohashi N, Rikihisa Y;
 XX DR WPI; 1999-254290/21.
 DR P-PSDB; AAY06962.

XX Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX Disclosure; Fig 22A; 55pp; English.
 CC The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX Sequence 843 BP; 283 A; 136 C; 150 G; 274 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 5,54e-34 Length: 843
 Score: 286.00 Matches: 97
 Percent Similarity: 51.09% Conservative: 44
 Best Local Similarity: 35.14% Mismatches: 115
 Query Match: 25.56% Indels: 20
 DB: 20 Gaps: 10
 US-09-846-808-1 (1-284) x AAX34762 (1-843)
 Qy 15 LeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn----- 32
 Db 43 TTAATGTACTATGCTCCAAGCATATCTTTTCTGTACTATACAGACGATAACACTGGT 102
 Qy 33 SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuLeuIle 52
 Db 103 AGCTTCTACATCAGTGGAAATATGTACCAAGTGTTCACATTTTGGTGTCTCAGCT 162
 Qy 53 LysGluSerAlaAlaAsnThrValGlnValPheGlyLeuLysLysAspLeuAsnAsp 72
 Db 163 AAAGAAGAAAGAACTCAACTGTTGGAGTTTGGATTAAACACATGATTGAATGGAGCT 222
 Qy 73 LeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIleLys---TyrAsnProTyrTyr 91
 Db 223 ACAATATCTACTTCTCCAGAAATATATTCACAGTTCAAAATATTCTGTTTAAATAC 282
 Qy 92 GluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsnLysAsnPhe 110
 Db 283 GAAACCAACCCATCTTTAGGTTTGGAGGACTATTGGTTTATTCATGGGT---GGCCCA 339
 Qy 111 ArgIleGluSerGluLeuSerTyrGluThrPheHisIleLys-----AsnAsnGlyTyr 128
 Db 340 AGAATAGAACTTGAAGTCTGTACGAGACATTCGATGTGAAAAATCAGAACATAATTAT 399
 Qy 129 LysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGlySerAsn 148
 Db 400 AAGAAC---GGCGCACACAGATCTGTCTTATCTCATCATAGTTCAGCAACAACATG 456
 Qy 149 AsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThrSerAla 168
 Db 457 TCCCTCCCAAGTAAACAAATTTGTTCTTAAAAATGAAGGTTAAATGACTTATCATTT 516
 Qy 169 LeuIleAsnValCysTyrAsp-----ValAspGlyLysHisAsnIleIleThrTyr 186
 Db 517 ATGATAATGCATGCTATGACATAATAATGGAAGAAATGCCCTTTTCA-----CCTTAT 570
 Qy 187 SerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrLysPhe 206
 Db 571 ATTTGTGCGGTGTTGGTACTGATGTTGTTCCATGTTTGAAGCTATATAATCCTAAAT 630
 Qy 207 SerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIle 226
 Db 631 TCTTACCAGGAAACCTAGGTTATAGTATAGTTTCAAGTTCAGAAGCCTCTGTTTATC 690
 Qy 227 GluGlyTyrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsnTyrPro 246
 Db 691 GGTGGACACTTTTCACAGATCATAGGTAATGAATTTAGAGACATCCCTGCTGCTCCT 750

QY 247 CysAspTyrProSerProThrProProAsnSerLysProHisValHisThrThrAlaLeu 266
 Db 751 AGTGA-----TCAATCTTCAGAAACCA-----TTTGAATA 786
 QY 267 AlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyIleLysPhe 282
 Db 787 GTAACACTAAATGTGTCACTTGTGTTAGACTTGGAGGAAGATT 834
 RESULT 22
 AAX34759
 ID AAX34759 standard; DNA; 867 BP.
 AC AAX34759;
 XX
 DT 05-JUL-1999 (first entry)
 XX
 DE DNA encoding P30 protein.
 XX
 KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
 KW detection; dog; ss.
 XX
 OS Ehrlichia canis.
 XX
 PN WO9913720-Al.
 XX
 PD 25-MAR-1999.
 XX
 PF 18-SEP-1998; 98WO-US19600.
 XX
 PR 19-SEP-1997; 97US-0059353.
 XX
 PA (OHIS) UNIV OHIO STATE.
 XX
 PI Ohashi N, Rikihisa Y;
 XX
 DR WPI; 1999-254290/21.
 DR P-PSDB; AAY06959.
 XX
 PT Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX
 PS Disclosure; Fig 19A; 55pp; English.
 CC
 CC The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the P30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX
 SQ Sequence 867 BP; 305 A; 150 C; 157 G; 255 T; 0 other;
 Alignment Scores:
 Pred. No.: 5.77e-34 Length: 867
 Score: 386.00 Matches: 99
 Percent Similarity: 50.70% Conservative: 46
 Best Local Similarity: 34.62% Mismatches: 113
 Query Match: 25.56% Indels: 28
 DB: 20 Gaps: 11
 US-09-846-808-1 (1-284) x AAX34759 (1-867)
 QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
 Db 37 ATATCACTAATGTCTTCTTACCTACCGTATCTTTTCTGAATCAATACATGAAGATAAT 96
 QY 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLys 49
 Db 97 ATAAATGGTAACCTTTTACATTTAGTGCAGAGTATATGCCAAGTGCCTCACACTTTGGCGTA 156
 QY 50 LeuLeuIleLysGlnSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp--- 68
 Db 157 TTTTCAGTTAAAGAGAGAGAAAAACACAACACTGGAGTTTCGGATTAAACAAAGATTGG 216

QY 69 -----LeuLeuAsnAspLeuLeuThrGly-----IleLysAspAsnThrAsnPhe 83
 Db 217 GACGGAGCAACAATAAAGGATGCAAGCAGCCACACACATAGACCCACCAATATTC 276
 QY 84 AsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIle 101
 Db 277 TCCATTTCAAATTTATTCATTTAAATATGAACAATCCATTTTATAGGTTTGCAGGAGCT 336
 QY 102 PheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThrPhe 121
 Db 337 ATTGGCTACTCAATGGGT---GGTCCAAGGTAGAGTTTGAAGTGTCTTACGAAATATT 393
 QY 122 HisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeu 139
 Db 394 GATGTAATAAACCAAGGTAACAGTTACAGAAC---GATGCTCACAATATATGCGCTTTA 450
 QY 140 AlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIle 159
 Db 451 TCAAGACACACCGGAGGTATGCCACAAGCCGGTCATCAAAATAAATTTGTCTTCTTAAAA 510
 QY 160 AsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeu 179
 Db 511 AATGAAGGATTTACTTGACATATCACTTATGATAAACGCATGTTATGATATAACAATCGAC 570
 QY 180 LysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeu 199
 Db 571 AGCATGCCATTTCTCCATATATATGCGAGGTATTTGGTAGTGACITAGTTTCGATGTTT 630
 QY 200 SerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSer 219
 Db 631 GAAACTACAAATCCTAAATTTCTTATCAAGGAAATTTAGGTGTAGTTACTCCATAAGC 690
 QY 220 ProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGlu 239
 Db 691 CCAGAAGCATCTCTTTTGTGGAGGACACTTTTCACAGAGCTTATAGGTAATGAATTTAAA 750
 QY 240 LysIleProValAsnTyrProCysAspTyrProSerProThrProProAsnSer----- 257
 Db 751 -----GACATTCCTGCAATAACTCTCTGCTGGAGCAACAGAA 786
 QY 258 ---LysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGly 276
 Db 787 ATTAAGGACACACAGTTTACAACA-----GTAACATTAAACATATGCCACTTCGGACTA 840
 QY 277 SerIleGlyIleLysPhe 282
 Db 841 GAGCTTGGAGGCGAGGTTT 858
 RESULT 23
 AAV07176
 ID AAV07176 standard; DNA; 864 BP.
 XX AAV07176;
 XX
 DT 14-SEP-1998 (first entry)
 XX
 DE Cowdria ruminantium MAP1 gene coding sequence.
 XX
 KW MAP1 gene; major antigenic protein 1; rickettsia; heartworm;
 KW DNA vaccine; ss.
 XX
 OS Cowdria ruminantium.
 XX
 PN WO9816554-Al.
 XX
 PD 23-APR-1998.
 XX
 PF 17-OCT-1997; 97WO-US19044.
 XX
 PR 17-OCT-1996; 96US-0733230.
 XX
 PA (UYFL) UNIV FLORIDA.

XX	Barbet AF, Burridge MJ, Ganta RR, Mahan SM, McGuire TC;
PI	Nyika A, Rurangirwa FR;
PI	
XX	WPI: 1998-251232/22.
DR	P-PSDB; AAW51088.
DR	
XX	Composition containing nucleic acid encoding rickettsial antigen -
PT	useful for, e.g. stimulating protective immune response in humans or
PT	animals
PT	
XX	Claim 4; Page 14-15; 39pp; English.
PS	
XX	This DNA molecule comprises the coding region of the major antigen
CC	protein 1 gene (WAPI) of Cowdria ruminantium, the causative agent
CC	of heartwater in domestic ruminants. It codes for a 287-amino acid
CC	WAPI protein (see AAW51088). A claimed composition comprises a
CC	nucleic acid (see AAW0176-82) encoding a polypeptide (see AAW51088-99)
CC	that elicits a protective immune response against a rickettsial
CC	pathogen. The nucleic acid is used, in human or veterinary
CC	medicine, in vaccines to protect against Rickettsia, Ehrlichia,
CC	Anaplasma and Cowdria species. The nucleic acids are also useful
CC	as probes to identify related sequences, e.g. for identification of
CC	organisms and for diagnosing infection. Use of nucleic acid
CC	vaccines avoids the problem of protein purification associated with
CC	protein-based vaccines. The nucleic acid does not replicate in the
CC	host but remains episomal and capable of expressing polypeptide for
CC	at least 19 mCh.
XX	
SO	Sequence 864 BP; 307 A; 139 C; 149 G; 269 T; 0 other;

Alignment Scores:			
Pred. No.:	1.le-33	Length:	864
Score:	383.50	Matches:	100
Percent Similarity:	50.17%	Conservative:	46
Best Local Similarity:	34.36%	Mismatches:	108
Query Match:	25.40%	Indels:	37
DB:	19	Gaps:	12
US-09-846-808-1 (1-284) x AAV07176 (1-864)			
Qy	13	ValMetLeuIleuPheThrProHisLeuAlaSerValLeu	----- 28
Db	37	ATATCATTAAGTGTCTATTTTACCTGGTGTGCTCTTTCGTGATATACAGGNAGACAC	96
Qy	29	AsnAspHisAsnSerMetYrValcGlyIleGlnTyrLysProAlaArgLlnHisLeuSer	48
Db	97	AACCCAGCAGGCAGTGTTTTATTAGCGCAAAATACATGCGCACTGCATCACATTTTGGT	156
Qy	49	LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp	68
Db	157	AAAATGTCATCAAGAAGATTCAAAANAATACTCAAAACGGTATTGTCTTAAANAAGAT	216
Qy	69	LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsnPhe	83
Db	217	-----TGGGATGGCGTTAAACACCACCATCAGATTCAGCAATACTAATTCT	261
Qy	84	AsnIle-----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe	98
Db	262	ACAAATTTTACTGAAAAAGACATATCTTTTCAGATATGAAAAACAATCGTTTTTAGTTTC	321
Qy	99	SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr	118
Db	322	GCRTGGAGCAATTGGGTACTCAATGANT---GGACCAAGATAGAGTTCGANGATATCCTAT	378
Qy	119	GluThrPheHisIleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHis	136
Db	379	GAACATTTTGATGTAAAAAACCTAGGTGGCAACTATAAAAC---AAGGCACACATGTAC	435
Qy	137	PheAlaLeuAlaLysGluIleSerGlyCysSerAsnAsnProAlaAsnAsnLysTyrVal	156
Db	436	TGTGCTTTAGATACAGCAGCACAAAATAGCACTAATCGCGCAGGATTAACTACATCTGTT	495

CC (MAP2) gene of rickettsial pathogens. The nucleic acid vaccines can be
 CC driven by the human cytomegalovirus (HCMV) enhancer-promoter. Cowdria
 CC ruminantium genes designated map 2, lhworf3, 4hworf1, 18hworf1 and
 CC 3dof3 may be used in therapeutic and diagnostic applications. The
 CC polypeptides are useful for detecting antibodies associated with
 CC infection by a rickettsial pathogen whilst the polynucleotides may be
 CC used to detect the presence of rickettsial nucleic acids.
 XX

SQ Sequence 864 BP; 307 A; 139 C; 149 G; 269 T; 0 other;

Alignment Scores:

Pred. No.: 1.1e-33 Length: 864
 Score: 383.50 Matches: 100
 Percent Similarity: 50.17% Conservative: 46
 Best Local Similarity: 34.36% Mismatches: 108
 Query Match: 25.40% Indels: 37
 DB: 21 Gaps: 12

US-09-846-808-1 (1-284) x AAC68699 (1-864)

QY 13 ValMetLeuLeuPheThrProHisIleSerLeuAlaSerValLeu----- 28
 Db 37 ATATCATTAGTGTGTCATTTTACCTGGTGTCTCTTTCTGATGTAATACAGGACAGC 96
 QY 29 AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
 Db 97 AACCCAGCAGCAGTGTTCATATTAGCGCAAAATACATGCACATGCATCATCTTTGGT 156
 QY 49 LysLeuLeuLeuLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68
 Db 157 AAATGTCAATCAAGAAGATTTCAAAATAACTCAACACGGTATTTGGTCTAAAAAAGAT 216
 QY 69 LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsnPhe 83
 Db 217 -----TGGGATGGCGTTAAACACCATCATCTTACATACTAATCT 261
 QY 84 AsnIle-----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
 Db 262 ACAATTTTACTGAAAAGACTATTCTTCAGATATGAAACATCCGTTTATAGTTTC 321
 QY 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
 Db 322 GCTGGCAATTTGGGTACTCAATGAAT---GGACCAAGATAGAGTTCGAAGTATCTAT 378
 QY 119 GluThrPheHisIleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHis 136
 Db 379 GAACTTTTGATGTAATAAACCTAGGTGGCACTATAAATAC---AACGCACATGTAC 435
 QY 137 PheAlaLeuAlaLysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
 Db 436 TGTGCTTTAGATACAGCAGCAGCAAAATAGCATAATGGCGCAGGATTAACATCATCTGTT 495
 QY 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
 Db 496 ATGTTAAACACGAAATTTAAACAATATATCATTAATGTTAAATGGGTGTATGATATC 555
 QY 177 -----AspGlyLeuLysHisAsnIleLeuThrTyrSerCysLeuGlyPheGlyValAsp 194
 Db 556 ATGCTTGTGATGATACCA-----GTTCTCCATATGATGTGCGAGTATGGCAGCTGAC 609
 QY 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
 Db 610 TTATGTGCAGTAATTAATGTACAAATCTAAATATTCTTATCAAGGAAAGTAGGCATA 669
 QY 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
 Db 670 AGTTACTCAATCAATCTGAAGCTTCTATCTTTATCTGCTGGACATTTCCATAGAGTTATA 729
 QY 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro 254
 Db 730 GGTAAATGAATTTAA-----GATATTGCTACTTAAAAATA 765
 QY 255 ProAsnSerLysProHisValHisThr-----AlaLeuAlaMetLeuSerIleGly 272

Db 766 TTTACTTCAAAAACAGGAATATCTAATCTGCTTTGCATCAGCAACACTGTGATGTTGT 825
 QY 273 TyrTyrGlyGlySerIleGlyIleLysPheIle 283
 Db 826 CACTTGTGTATAGAAATTTGGAGGAAGGTTTGTA 858
 RESULT 25
 AAS07575
 ID AAS07575 standard; DNA; 864 BP.
 XX
 AC AAS07575;
 XX
 DT 23-OCT-2001 (first entry)
 XX
 DE DNA encoding major antigenic protein 1 (MAP1).
 XX
 KW Major antigenic protein 1; MAP1; vaccine; immunogenic; rickettsia;
 XX infection; heartwater; diagnostic; ds.
 OS Cowdria ruminantium.
 FH Key Location/Qualifiers
 FT CDS 1..864
 FT /*tag= a
 FT /product= "Major antigenic protein 1 (MAP1)"
 XX
 PN US6251872-B1.
 XX
 PD 26-JUN-2001.
 XX
 PF 17-OCT-1997; 97US-0953326.
 XX
 PR 17-OCT-1996; 96US-0733230.
 XX
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Barbet AF, Ganta RR, McGuire TC, Burridge MJ, Nyika A;
 PI Rurangirwa FR, Mahan SM, Bowie MV, Alleman AR;
 XX
 DR WPI; 2001-424487/45.
 DR P-PSDB; AAU04192.
 XX
 PT New MAP2 genes and polypeptides useful as vaccines for conferring
 PT immunity to human and animal rickettsial diseases, e.g. heartwater, or
 PT as molecular markers in nucleic acid analysis procedures -
 XX
 PS Example 1; Fig 1; 30pp; English.
 XX

CC The sequence represents the coding sequence of major antigenic protein 1
 CC (MAP1) isolated from Cowdria ruminantium. The MAP polynucleotides and
 CC polypeptides are useful as vaccines for conferring immunity to rickettsia
 CC infection, including Cowdria ruminantium causing heartwater. The MAP
 CC polynucleotides may be used as molecular markers in nucleic acid
 CC analysis procedures, and to produce the MAP polypeptides, which may
 CC be used to raise antibodies that are reactive with the polypeptides.
 CC The nucleic acids may further be used as probes to identify
 CC complementary sequences within other nucleic acid molecules or genomes,
 CC where such probes can be applied to identify or distinguish infectious
 CC strains of organisms in diagnostic procedures or in rickettsial
 CC research where identification of particular organisms or strains is
 CC needed.
 XX
 SQ Sequence 864 BP; 307 A; 139 C; 149 G; 269 T; 0 other;

Alignment Scores:

Pred. No.: 1.1e-33 Length: 864
 Score: 383.50 Matches: 100
 Percent Similarity: 50.17% Conservative: 46
 Best Local Similarity: 34.36% Mismatches: 108
 Query Match: 25.40% Indels: 37
 DB: 22 Gaps: 12

ID AAV07177 standard; DNA; 842 BP.
 AC AAV07177;
 DT 14-SEP-1998 (first entry)
 XX Ehrlichia chaffeensis MAP1 gene coding sequence.
 DE
 KW MAP1 gene; major antigenic protein 1; rickettsia; DNA vaccine; ss.
 XX Ehrlichia chaffeensis.
 OS
 PN WO9816554-AL.
 PD 23-APR-1998.
 XX
 PF 17-OCT-1997; 97WO-US19044.
 XX
 PR 17-OCT-1996; 96US-0733230.
 XX
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Barbet AF, Burr ridge MJ, Ganta RR, Mahan SM, McGuire TC;
 PI Nyika A, Rurangirwa FR;
 XX
 DR WPI; 1998-251232/22.
 DR P-PSDB; AAW51089.
 XX
 CC Composition containing nucleic acid encoding rickettsial antigen -
 PT useful for, e.g. stimulating protective immune response in humans or
 PT animals
 XX
 PS Claim 4; Page 17-18; 39pp; English.
 XX
 CC This DNA molecule comprises the coding region of the major antigen
 CC protein 1 gene (MAP1) of Ehrlichia chaffeensis. It codes for a 280
 CC amino acid MAP1 protein (see AAW51089). A claimed composition
 CC comprises a nucleic acid (see AAV07176-82) encoding a polypeptide
 CC (see AAW51088-99) that elicits a protective immune response against a
 CC rickettsial pathogen. The nucleic acid is used, in human or
 CC veterinary medicine, in vaccines to protect against Rickettsia,
 CC Ehrlichia, Anaplasma and Cowdria species. The nucleic acids are
 CC also useful as probes to identify related sequences, e.g. for
 CC identification of organisms and for diagnosing infection. Use of
 CC nucleic acid vaccines avoids the problem of protein purification
 CC associated with protein-based vaccines. The nucleic acid does not
 CC replicate in the host but remains episomal and capable of
 CC expressing polypeptide for at least 19 mth.
 XX
 SQ Sequence 842 BP; 282 A; 143 C; 166 G; 251 T; 0 other;

 Alignment Scores:
 Pred. No.: 3,56e-32 Length: 842
 Score: 370.00 Matches: 95
 Percent Similarity: 49.12% Conservative: 45
 Best Local Similarity: 33.33% Mismatches: 113
 Query Match: 24.50% Indels: 32
 DB: 19 Gaps: 9

 US-09-846-808-1 (1-284) x AAV07177 (1-842)

 QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAla-----SerValLeuAsn 29
 DB 37 ATCATTAATATCTCTCTTACTCGAGTATCATTTTCGACCCCAAGCGAGTAGTGGTC 96
 QY 30 AspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLys 49
 DB 97 ATTAACGGTAATTTCTACATCAGTGGAAAAATACGATGCCAAGGCTTCGCATTTCGAGTA 156
 QY 50 LeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeu 69
 DB 157 TTCTCTGCTAAGGAAGAAAGAAATACACAGTGGAGTGTGGAGCTGAAGCAAAATGG 216

QY 70 LeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsn 88
 DB 217 GACGGAGCGCAATATCCAACTCTCCCAACAGTATTCACCTGCTCAAAATATTCA 276
 QY 89 ProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrAsn 107
 DB 277 TTTAAATATGAACAACACCCGTTTGTAGGTTTTCAGGAGCTATTGGTTACTCAATGGAT 336
 QY 108 LysAsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn----- 125
 DB 337 ---GGTCCAAGAATAGAGCTTGAAGTATCTTATGAACATTTGATGATAAAATCAAGGT 393
 QY 126 AsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGly 145
 DB 394 AACAATTATAAGAAAT---GAAGCACATAGATATGTGCTCTATCCCATAC---TCAGCA 447
 QY 146 GlySerAsnAsnProAlaAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeu 165
 DB 448 GCAGACATGAGTAGTCAAGTAAATTTGTCTTCTTAAAAAATGAAGGATTACTTGGAC 507
 QY 166 ThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHisAsnIleIleThr 185
 DB 508 ATATCATTTATGCTGAACGCATGCTATGAGTAGTAGCGAAGCATACCTTTTCTCCT 567
 QY 186 TyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrLys 205
 DB 568 TATATATGCGCAGGTATCGTACTGATTTAGTATCCATGTTTGAAGCTACAAATCCTAAA 627
 QY 206 PheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPhe 225
 DB 628 ATTCTCTACCAAGAAAGTATAGGTTTAAAGCTACTATTAAGCCCAAGGCTTCGTGTTT 687
 QY 226 IleGluGlyTyrThrHisGlyLeuPheGlyLysPheGlyLysIleProValAsnTyr 245
 DB 688 ATGTGGGCGACATTTCATAGGTAATAGGGAACCAATTTAGACATATTCCTACTATAATA 747
 QY 246 ProCys-----AspTyrProSerProThrProProAsnSer 257
 DB 748 CCTACTGGATCAACACTTGCAGGAAAAAGAACTACCT----- 786
 QY 258 LysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySer 277
 DB 787 -----GCATAGTAATACTGGATGTATGCCACTTTGGAATAGAA 825
 QY 278 IleGlyIleLysPhe 282
 DB 826 ATGGGAGGAGGTTT 840
 RESULT 34
 AAC68700
 ID AAC68700 standard; DNA; 842 BP.
 XX
 AC AAC68700;
 XX
 DT 02-MAR-2001 (first entry)
 XX
 DE Ehrlichia chaffeensis MAP1 gene.
 XX
 KW Ehrlichia chaffeensis; MAP1; major antigenic protein 1; antirickettsial;
 KW vaccine; gene therapy; Rickettsia; Cowdria; Anaplasma; map2; lhworf3;
 KW 4hworfl; 18hworfl; 3gdorf3; ds.
 OS
 XX Ehrlichia chaffeensis.
 PN WO2000065063-A2.
 XX
 PD 02-NOV-2000.
 XX
 PF 21-APR-2000; 2000WO-US10886.
 XX
 PR 22-APR-1999; 99US-0130725.
 XX
 PA (UYFL) UNIV FLORIDA.

CC polypeptides are useful as vaccines for conferring immunity to rickettsia
 CC infection, including Cowdria ruminantium causing heartwater. The MAP
 CC polynucleotides may be used as molecular markers in nucleic acid
 CC analysis procedures, and to produce the MAP polypeptides, which may
 CC be used to raise antibodies that are reactive with the polypeptides.
 CC The nucleic acids may further be used as probes to identify
 CC complementary sequences within other nucleic acid molecules or genomes,
 CC where such probes can be applied to identify or distinguish infectious
 CC strains of organisms in diagnostic procedures or in rickettsial
 CC research where identification of particular organisms or strains is
 CC needed.

XX SQ Sequence 842 BP; 282 A; 143 C; 166 G; 251 T; 0 other;

Alignment Scores:

Pred. No.: 3,56e-32 Length: 842
 Score: 370.00 Matches: 95
 Percent Similarity: 49.12% Conservative: 45
 Best Local Similarity: 33.33% Mismatches: 113
 Query Match: 24.50% Indels: 32
 DB: 22 Gaps: 9

US-09-846-808-1 (1-284) x AAS07576 (1-842)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAla-----SerValLeuAsn 29
 DB ATCATTAATATCTCTCTTACCTGGAGTATCATTTTCGACCCCAAGCAGGTAGTGTC 96
 QY 30 AspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLys 49
 DB ATTAACGGTAATTTTACATCAGTGGAAATACGATGCCAAGGCTTCGCATTTTGGAGTA 156
 QY 50 LeuLeuIleLeuSerAlaAlaAsnThrValGluValPheGlyLeuLysAspLeu 69
 DB TTCTCTGCTAAGCAAGAAATACACAGTGGNGTGTGGACTGAAGCAAAATGG 216
 QY 70 LeuAsnAspLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsn 88
 DB GACGGAAGGCAATATCCAACTCTCCCAACAGATGATTCACCTGTCTCAAAATTATTC 276
 QY 89 ProTyrTyrGluAsnArg---LeuGlyPheSerGlyLePheGlyTyrTyrAsn 107
 DB TTTAATATGAACAACCCGTTTGTAGGTTTTCGAGGAGCTATTGGTTACTCAATGGAT 336
 QY 108 LysAsnPheArgIleGlySerGluLeuSerTyrGluThrPheHisIleLysAsn----- 125
 DB ---GGTCCCAAGATAGAGCTTGAAGTATCTTATGAACATTTGATGATAAAATCAAGCT 393
 QY 126 AsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluLeSerGly 145
 DB AACAAATTATAAAGAT---GAAGCACATAGATATTGTCTCTATCCATAAC---TCAGCA 447
 QY 146 GlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeu 165
 DB GCAGACATCAGTAGTCAAGTAAATATTTTGTCTTCTAAATAAATGAAGGATTACTTGAC 507
 QY 166 ThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHisAsnIleThr 185
 DB ATATCATTTATGCTGAACCATGCTATGACGTAGTAGGCGAAGGCATACCTTTTCTCCT 567
 QY 186 TyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrLys 205
 DB TATATATGCCAGGTATCGTACTGATTTAGTATCCATGTTTGAAGCTACAAATCCTAAA 627
 QY 206 PheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPhe 225
 DB ATTCTTACCAAGGAAGTTAGTTTAAGCTACTCTATAGGCCCAAGACCTCTGTGTTT 587
 QY 226 IleGluGlyTyrTyrHisGlyPhePheGlyLysLysPheGluLysIleProValAsnTyr 245
 DB ATTTGGTGGCCTTCTTCAATAGGTAATAGGGAACGAATTTAGAGATATTCCTACTATAATA 747
 QY 246 ProCys-----AspTyrProSerProThrProAsnSer 257

DB 748 CCTACTGGATCAACACTTTCGAGAAAGAACTACCT----- 786
 QY 258 LysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySer 277
 DB 787 -----GCAATAGTAATACTGGATGATGCCACITTTGGAATAGAA 825
 QY 278 IleGlyIleLysPhe 282
 DB 826 ATGGGAGGAAGGTTT 840
 RESULT 36
 AAX34761
 ID AAX34761 standard; DNA; 924 BP.
 XX AAX34761;
 XX 05-JUL-1999 (first entry)
 DT DNA encoding p30-1 protein.
 DE Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; p30;
 KW detection; dog; ss.
 XX Ehrlichia canis.
 OS WO9913720-Al.
 PN 25-MAR-1999.
 PD 18-SEP-1998; 98WO-US19600.
 PF 19-SEP-1997; 97US-0059353.
 PR (OHIS) UNIV OHIO STATE.
 PA Ohashi N, Rikihisa Y;
 PI WPI; 1999-254290/21.
 DR P-PSDB; AAY06961.
 XX Novel outer membrane proteins from Ehrlichia chaffeensis and
 PT Ehrlichia canis
 XX Disclosure; Fig 21A; 55pp; English.
 CC The invention provides isolated outer membrane proteins (OMP) from
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
 CC of the OMP family and consist of proteins OMP-1, -1(B to Z) shown
 CC in AAY06943-958. The E. canis proteins form part of the p30 family and
 CC consist of proteins shown in AAY06959-970. The proteins and genes are
 CC used to detect E. chaffeensis in patients and E. canis in dogs.
 XX Sequence 924 BP; 301 A; 158 C; 160 G; 305 T; 0 other;
 SQ

Alignment Scores:

Pred. No.: 2,52e-31 Length: 924
 Score: 363.00 Matches: 94
 Percent Similarity: 50.18% Conservative: 45
 Best Local Similarity: 33.94% Mismatches: 118
 Query Match: 24.04% Indels: 20
 DB: 20 Gaps: 9

US-09-846-808-1 (1-284) x AAX34761 (1-924)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
 DB 124 ATATCATTAATGACTCTCTTATCCCAAGCATATCTTTTCGTACTATCAAGATGGTAAC 183
 QY 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLys 49
 DB 184 ATGGGTGGTAACTTCTATATTAGTGGGAAGATGTACCAAGTCTCTCACATTTTGTAGC 243

```
Qy 50 LeuLeuLeuLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeu 69
Db 244 TTCACAGTAAGAAGAAAGCAATCACTGTTGGAGTTTTTGGATTAAACATGATTGG 303
Qy 70 LeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsn 88
Db 304 GATGGAAGTCCAATACTTAAGAATAAA---CACGCTGACTTTACTGTTCACAACTATTTCG 360
Qy 89 ProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsn 107
Db 361 TTCAGATACGAGAACATCCATTTCTAGGGTTTCAGGAGCTATCGGTACTCAATGGGT 420
Qy 108 LysAsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLys-----Asn 125
Db 421 ---GGCCCAAGATAGATTGCAATATCTTATGAGCATTCGAGCTAAAAGTCCCTAAT 477
Qy 126 AsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGly 145
Db 478 ATCAATTATCAAAAT---GACGCGCACAGGTACTGCGCTCTATCTCATCACATCGGCA 534
Qy 146 GlySerAsnAsnProAlaAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeu 165
Db 535 GCCATGGAA-----GCTGATAAATTTGCTCTTAAATAAACGAAGGGTTAATTGAC 585
Qy 166 ThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHisAsnIleIleThr 185
Db 586 ATATCACTTGCAATAATGCATGTATGATATATAATATGACAAAGTACCTGTTCTCCT 645
Qy 186 TyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrLys 205
Db 646 TATATATCGCAGGTATTGGTACTGATTTGATTCTATGTTTGAAGCTACAAAGTCTCTAAA 705
Qy 206 PheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPhe 225
Db 706 ATTCTTACCAGGAATACTGGCATTTAGTTACTCTATTAATCGGAACCTCTGTTTC 765
Qy 226 IleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsnTyr 245
Db 766 ATCGTGGGCATTTCCACAGCATATAGTAAATGAGTTTAGAGATATTCCTGCAATAGTA 825
Qy 246 ProCysAspTyrProSerProThrProProAsnSerLysProHisValHisThrAla 265
Db 826 CCTAGTAACCTCAACTACAAATAAGTGGACCAAA-----TTTGCA 864
Qy 266 LeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyIleLysPhe 282
Db 865 ACAGTAACACTAAATATGTGTGCTTTGGTTTAGAACTTTGGAGGAAGATT 915

RESULT 37
AAD01292
ID AAD01292 standard; DNA; 1607 BP.
XX
AC AAD01292;
XX
DT 12-OCT-2000 (first entry)
XX
DE Ehrlichia canis immunoreactive protein Eca28-1 DNA.
XX
KW Homologous 28-kDa protein gene; Eca28-1; immunoreactive; vaccine;
p28 gene; polymorphic multiple gene family; immunoprotective antigen;
antibacterial; canine ehrlichiosis; canine tropical pancytopenia;
tick-borne rickettsial disease; serodiagnosis; ds.
XX
OS Ehrlichia canis.
XX
FH Key Location/Qualifiers
CDS 146..982
FT /*tag= a
FT /product= "Eca28-1 protein (30-kDa)"
FT sig_peptide 146..214
FT /*tag= b
FT mat_peptide 215..979
FT /*tag= c
```

```
FT /product= "Mature Eca28-1 protein (28-kDa)"
XX
PN WO200032745-A2.
XX
PD 08-JUN-2000.
XX
PF 24-NOV-1999; 99WO-US28075.
XX
PR 30-NOV-1998; 98US-0201458.
PR 03-MAR-1999; 99US-0261358.
XX
PA (RERE-) RES DEV FOUND.
XX
PI Walker DH, Yu X, McBride JW;
XX
DR WPI; 2000-412298/35.
DR P-PSDB; AAY71477.
XX
PT Ehrlichia canis antigens useful for vaccinating against canine
ehrichiosis in dogs
XX
PS Claim 5; Fig 1; 86pp; English.
XX
CC The patent relates to homologous 28-kilodalton (kDa) protein genes of
Ehrlichia canis, designated Eca28SA1, Eca28SA2, Eca28SA3, Eca28-1 and
Eca28-2. These genes are members of a polymorphic multiple gene family
and contained in a single locus of 5.592 kb. The 28-kDa proteins are
immunoreactive with anti-E. canis serum hence are important
immunoprotective antigens. The protein is useful for vaccinating
against E. canis infections such as canine ehrlichiosis in dogs.
CC Canine ehrlichiosis, also known as canine tropical pancytopenia, is a
tick-borne rickettsial disease of dogs. Eca28-1 is conserved amongst
different strains of E. canis and hence useful for serodiagnosis of
canine ehrlichiosis. The present sequence is a DNA encoding E. canis
Eca28-1 30-kDa protein which is post-translationally modified to a
mature 28-kDa protein by cleavage of N-terminal signal sequence.
XX
SQ Sequence 1607 BP; 543 A; 268 C; 250 G; 540 T; 6 other;

Alignment Scores:
Pred. NO.: 5.59e-31 Length: 1607
Score: 363.00 Matches: 94
Percent Similarity: 50.18% Conservative: 45
Best Local Similarity: 33.94% Mismatches: 118
Query Match: 24.04% Indels: 20
DB: 21 Gaps: 9

US-09-846-808-1 (1-284) x AAD01292 (1-1607)
Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
Db 182 ATATCATTAATGTTACTCTATTCCACGACATATCTTTCTGTACTACTATACAGATGGTAAC 241
Qy 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLys 49
Db 242 ATGGGTGGTAACCTCTATATTAGTGGAAAGTATGTACCAAGTGTCTCACATTTGGTAGC 301
Qy 50 LeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeu 69
Db 302 TTCTCAGCTAAGAAGAAAGCAATCAACTGTTGGAGTTTTTGGATTAAACATGATTGG 361
Qy 70 LeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsn 88
Db 362 GATGGAAGTCCAATACTTAAGAATAAA---CACGCTGACTTTTACTTCCAAACTATTTCG 418
Qy 89 ProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsn 107
Db 419 TTCAGATACGAGACAATCCATTTCTAGGGTTTCGAGGAGCTATCGGTACTCAATGGGT 478
Qy 108 LysAsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLys-----Asn 125
Db 479 ---GGCCCAAGAATAAGATTGCAATATCTTATGAAGCATTCGAGCTAAAAAGTCTCTAAT 535
```


QY 126 AsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGly 145
 Db 536 ATCAATATCAAAAT---GACGCGCAGGTAAGTGGCTCTATCTCATCACACATCGGCA 592
 QY 146 GlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeu 165
 Db 593 GCATGGAA-----GCTGATAAATTTGCTCTTTAAACAAAGGAGGTTAATTGAC 643
 QY 166 ThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHisAsnIleThr 185
 Db 644 ATATCACTTGCATAAATCATGTTATGATATAATAATGACAAAGTACTGTTCTCCT 703
 QY 186 TyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrLys 205
 Db 704 TATATATGCGCAGGTAATTCGTAATTCGTAATTCGTAATTCGTAATTCGTAATTCGTA 763
 QY 206 PheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPhe 225
 Db 764 ATTTCTTACCAAGGAAACTGGGCATTTAGTTACTTATTAATCCGGAACCTCTGTTTC 823
 QY 226 IleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGlyLysProValAsnTyr 245
 Db 824 ATCGGTGGCATTCCACAGGATCATAGTAATGAGTTAGAGATATCTCTCAATAGTA 883
 QY 246 ProCysAspTyrProSerProThrProProAsnSerLysProHisValHisThrAla 265
 Db 884 CCTAGTAACCTCACTACAATAAGTGGACCAAA-----TTTGCA 922
 QY 266 LeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyIleLysPhe 282
 Db 923 ACAGTAACATAAATGTGTCTACTTGTCTTGTAGAACCTGGGAGAGATTT 973
 RESULT 38
 AAV07180
 ID AAV07180 standard; DNA; 1570 BP.
 XX AC AAV07180;
 XX DT 14-SEP-1998 (first entry)
 XX DE Ehrlichia canis VSA genomic locus.
 XX KW MAPI homologue; variable surface antigen; VSA1; VSA2; rickettsia;
 KW DNA vaccine; ss.
 XX OS Ehrlichia canis.
 XX FH Key Location/Qualifiers
 FT RBS 1..5
 FT /*tag= a
 FT CDS 11..874
 FT /*tag= b
 FT /*note= "VSA1 gene"
 FT terminator 1015..1022
 FT /*tag= c
 FT terminator 1027..1034
 FT /*tag= d
 FT -35_signal 1081..1086
 FT /*tag= e
 FT exon 1101..1106
 FT /*tag= f
 FT RBS 1161..1165
 FT /*tag= g
 FT CDS 1171..1560
 FT /*tag= h
 FT /*note= "truncated VSA2 gene"
 XX PN W09816554-A1.
 XX PD 23-APR-1998.
 XX PF 17-OCT-1997; 97WO-US19044.
 XX

PR 17-OCT-1996; 960S-0733230.
 XX (UYFL) UNIV FLORIDA.
 XX Barbet AF, Burr ridge MJ, Ganta RR, Mahan SM, McGuire TC;
 PI Nyika A, Rurangirwa FR;
 XX WPI; 1998-251232/22.
 DR P-PSDB; AAW51096-97.
 PT Composition containing nucleic acid encoding rickettsial antigen -
 PT useful for, e.g. stimulating protective immune response in humans or
 PT animals
 XX PS Claim 4; Fig 2C; 39pp; English.
 CC This is the DNA sequence of a 1.5 genomic locus of Ehrlichia canis
 CC that was obtained using a PCR cloning strategy based on identifying
 CC genes homologous to the major antigenic protein MAPI (see AAW51088)
 CC of Cowdria ruminantium. It includes 2 very similar but
 CC non-identical open reading frames (ORFs), of which ORF2 is a
 CC partial gene. Due to their similarity to MAPI surface antigen
 CC genes of C. ruminantium, the E. canis ORFs are designated variable
 CC surface antigen (VSA) genes 1-2. A claimed composition comprises a
 CC nucleic acid (see AAV07176-82) encoding a polypeptide (see AAW51088-99)
 CC that elicits a protective immune response against a rickettsial
 CC pathogen. The nucleic acid is used, in human or veterinary
 CC medicine, in vaccines to protect against Rickettsia, Ehrlichia,
 CC Anaplasma and Cowdria species. The nucleic acids are also useful
 CC as probes to identify related sequences, e.g. for identification of
 CC organisms and for diagnosing infection. Use of nucleic acid
 CC vaccines avoids the problem of protein purification associated with
 CC protein-based vaccines. The nucleic acid does not replicate in the
 CC host but remains episomal and capable of expressing polypeptide for
 CC at least 19 mth.
 XX SQ Sequence 1570 BP; 565 A; 223 C; 251 G; 531 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,12e-28 Length: 1570
 Score: 342.50 Matches: 95
 Percent Similarity: 45.70% Conservative: 43
 Best Local Similarity: 31.46% Mismatches: 129
 Query Match: 22.68% Indels: 35
 DB: 19 Gaps: 9
 US-09-846-808-1 (1-284) x AAV07180 (1-1570)
 QY 1 MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuLeu----- 17
 Db 5 GTAAATATGAAATATAAATAAATCTTTACAGTAACCTGCATTAGTATTATTAACCTCTTT 64
 QY 18 -----PheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsnSerMetTyr 35
 Db 65 ACACATTTTATACCTTTTATAGTCAGCAGCGTCACATTTTGGAAATTTTTCAGCTAAGAAGAA 121
 QY 36 ValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSer 55
 Db 122 ATTAGTGGAATAATATATGCGCACAGCGTCACATTTTGGAAATTTTTCAGCTAAGAAGAA 181
 QY 56 AlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeu----- 70
 Db 182 CAAAGTTTTACTAAGGTATTAGTTGGTTAGATCAACAGCATTTATACATAATATTATAAAC 241
 QY 71 ---AsnAspLeuLeuThrGlyIleLys---AspAsnThrAsnPheAsnIleLysTyrAsn 88
 Db 242 AATAATGATACAGCAAGAGCTCTTAAGGTTCAAAATTTTCAATTAATAACAAATAAAC 301
 QY 89 ProTyrTyrGluAsnAsnArgLeuGlyPheSerGlyIlePheGlyTyrTyrAsnLys 108
 Db 302 CCATTT-----CTAGGATTTGCAGGAGCTATTGGTTATTCAATAGGC--- 343
 QY 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsnAsnGlyTyr 128

```

Db 344 AATCAAGATAGACTAGAGTATACATGAATATATTGTACTATAAAACCCAGGAAC 403
Qy 129 Lysargile---AspCysGLuLYsHisPheAlaLeuAlaLYsGLuIle----- 143
Db 404 AATTATTAAATGACTCTCACAAATATTGCCCTTATCTCATCGAAGTCACATGCACT 463
Qy 144 -----SerGlyGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuLea 160
Db 464 GATGAAATACGGAGATTGGTACACTGCAAAACTGATAAGTTGTACTTCTGAAAAAT 523
Qy 161 AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeu 180
Db 524 GAAGTTTACTGACGCTCTATTATGTTAAACGATGTTATGACATAACAACGAAAAA 583
Qy 181 HisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSer 200
Db 584 ATGCCCTTTTACCTTATATATGTCAGGTATTGGTACTGATCTCATATCTGTTGAG 643
Qy 201 LysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerPro 220
Db 644 ACAACACAAAAAATATCTTATCAAGGAAAGTTAGGTTTAAACTATATAAACTCA 703
Qy 221 GlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGlu 240
Db 704 AGAGTTTCTGTTTTCAGAGTGGGCACCTTCATTAAGGTAATAGGTAATGAATTTAAAGGT 763
Qy 241 IleProValAsnTyrProCysAspTyrProSerProThrProProAsnSerLysProHis 260
Db 764 ATT-----CCTACTCTATTACCTGATGGTCAACAACTTAAA 799
Qy 261 ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrGlyGlySerIleGlyIle 280
Db 800 GTACAACAGTGTCAACAGTAGTAACATTAGATGTCGCCATTTTCGGTTAGAGATTGGAAGT 859
Qy 281 Lysphe 282
Db 860 AGATTT 865

RESULT 39
ID AAS07583 standard; DNA; 1570 BP.
XX
AC AAS07583;
XX
XX 23-OCT-2001 (first entry)
XX
XX DNA encoding variable surface antigens 1-2 (VSA1-2) from E. canis.
XX
XX Major antigenic protein; MAP; vaccine; immunogenic; rickettsia;
KW infection; heartwater; diagnostic; variable surface antigen; VSA; ds.
XX
XX Ehrlichia canis.
XX
XX Key Location/Qualifiers
FT RBS 1..5
FT /*tag= a
FT /note= "Ribosome binding site of VSA1 gene"
FT CDS 11..874
FT /*tag= b
FT /product= "Variable surface antigen 1 (VSA1)"
FT terminator 1015..1020
FT /*tag= c
FT /note= "Transcription terminator of VSA1 gene"
FT terminator 1027..1034
FT /*tag= d
FT /note= "Transcription terminator of VSA1 gene"
FT -35_signal 1081..1086
FT /*tag= e
FT /note= "-35 region of VSA2 gene"
FT -10_signal 1101..1106
FT /*tag= f
FT /note= "-10 region of VSA2 gene"
FT
```

```

FT RBS 1151..1155
FT /*tag= g
FT /note= "Ribosome binding site of VSA2 gene"
FT CDS 1161..1169
FT /*tag= h
FT /product= "variable surface antigen 2 (VSA2)"
FT /partial
FT /note= "No stop codon"
FN US6251872-B1.
XX
XX 26-JUN-2001.
XX
XX 17-OCT-1997; 97US-09533326.
XX
XX 17-OCT-1996; 96US-0733230.
XX
XX (UYFL ) UNIV FLORIDA.
XX
XX Barbet AF, Ganta RR, McGuire TC, Burridge MJ, Nyika A;
XX Rurangirwa FR, Mahan SM, Bowie MV, Alleman AR;
XX
XX WPI: 2001-424487/45.
XX
XX P-PSDB; AAU04200, AAU04201.
XX
XX New MAP2 genes and polypeptides useful as vaccines for conferring
XX immunity to human and animal rickettsial diseases, e.g. heartwater, or
XX as molecular markers in nucleic acid analysis procedures -
XX
XX Example 3; Fig 2C; 30pp; English.
XX
XX The sequence represents the coding sequence of variable surface antigen
XX (VSA) gene locus encoding VSA1-2 proteins of Ehrlichia canis, which
XX have similarity to major antigen proteins (MAP). The MAP polynucleotides
XX and polypeptides are useful as vaccines for conferring immunity to
XX rickettsia infection, including Cowdria ruminantium causing heartwater.
XX The MAP polynucleotides may be used as molecular markers in nucleic acid
XX analysis procedures, and to produce the MAP polypeptides, which may
XX be used to raise antibodies that are reactive with the polypeptides.
XX The nucleic acids may further be used as probes to identify
XX complementary sequences within other nucleic acid molecules or genomes,
XX where such probes can be applied to identify or distinguish infectious
XX strains of organisms in diagnostic procedures or in rickettsial
XX research where identification of particular organisms or strains is
XX needed.
XX
XX Sequence 1570 BP; 565 A; 223 C; 251 G; 531 T; 0 other;
```

Alignment Scores:

Pred. No.: 1.12e-28 Length: 1570
Score: 342.50 Matches: 95
Percent Similarity: 45.70% Conservative: 43
Best Local Similarity: 31.46% Mismatches: 129
Query Match: 22.68% Indels: 35
DB: 22 Gaps: 9

US-09-846-808-1 (1-284) x AAS07583 (1-1570)

```

Qy 1 MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuIleLeu----- 17
Db 5 GTAAATATGAAATATAAAAAAATTTTACAGTAACTGCATTAGTATTATTACTTCCTTT 64
Qy 18 -----PheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsnSerMetTyr 35
Db 65 ACACATTTTATACCTTTTATAGTCCACGCGTCCAGTACAAATTCACACAC---TTCTAC 121
Qy 36 ValGlyIleGlnTyrLysProAlaArgGlnHisSerLysLeuLeuIleLysGluSer 55
Db 122 ATTAGTGGAATAATATATGCCAACACGCGTCACATTTTGGAAATTTTTCAGCTAAAGAAGAA 181
Qy 56 AlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeu----- 70
Db 182 CAAAGTTTACTAAGGTATTAGTGGGTTAGATCAACGATTATCATCAATAATATTATAAAC 241
```

```
QY 71 ---AsnAspLeuLeuThrGlyIleLys---AspAsnThrAsnPheAsnIleLysTyrAsn 88
DB 242 AATAATGATACAGCAAGAGCTCTAAGGTTCAAAATATTATTATTAATAACAAAATAAC 301
QY 89 ProTyrTyrGluAsnAsnArgLeuGlyPheSerGlyIlePheGlyTyrTyrAsnLys 108
DB 302 CCATTT-----CTAGGATTTGCAGGAGCTATTGGTTATTCAATAGGC--- 343
QY 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsnAsnGlyTyr 128
DB 344 AATTCAGAAATAGACTAGAGTATACATGAAATATTGATCATTAACCCAGGAAAC 403
QY 129 LysArgIle---AspCysGluLysHisPheAlaLeuAlaLysGluIle----- 143
DB 404 AATTATTAAATGACTCTCACAATAATTGCGCTTATCTCATGGAAGTACATATGCACT 463
QY 144 -----SerGlyCysSerAsnProAlaAsnAsnLysTyrValThrLeuIleAsn 160
DB 464 GATGGAATAGCGGAGATTGGTACACTGCAGGAGTATGGTACTCATATCTATGTTGAG 523
QY 161 AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLys 180
DB 524 GAAGGTTTACTTCACCTCTCATTTATGTTAAACGCATGTTATGACATACAACTGAAATA 583
QY 181 HisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSer 200
DB 584 ATGCCCTTTTTCACCTTATATATGTCAGGATTTGGTACTGATCTCATATCTATGTTGAG 643
QY 201 LysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerPro 220
DB 644 ACAACAAACAAACAAATATCTTATCAAGGAAAGTTAGGTTTAAACTATATATAACTCA 703
QY 221 GlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGlyLys 240
DB 704 AGAGTTTCTGTTTTCAGGTCAGGTCGACCTTTCATAGGTAATAGGTAATGAATTAAGCT 763
QY 241 IleProValAsnTyrProCysAspTyrProSerProThrProAsnSerLysProHis 260
DB 764 ATT-----CCTACTCTATTACCTGATGGATCAACATTAATAA 799
QY 261 ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyCysIleGlyIle 280
DB 800 GTACAACAGCTCTGCAACAGATACATTAGATGTGTGCCATTTCCGGGTTAGAGATTGGAAGT 859
QY 281 LysPhe 282
DB 860 AGATTT 865
RESULT 40
AA34764
ID AAX34764 standard; DNA; 831 BP.
XX
AC AAX34764;
XX
DT 05-JUL-1999 (first entry)
XX
DE DNA encoding p30-4 protein.
XX
KW Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; p30;
KW detection; dog; ss.
XX
OS Ehrlichia canis.
XX
PN W09913720-A1.
XX
PD 25-MAR-1999.
XX
PF 18-SEP-1998; 98WO-US19600.
XX
PR 19-SEP-1997; 97US-0059353.
XX
PA (OHIS ) UNIV OHIO STATE.
```

```
XX
PI Ohashi N, Rikihisa Y;
XX
DR WPI: 1999-254290/21.
DR P-PSDB; AAY06964.
XX
PT Novel outer membrane proteins from Ehrlichia chaffeensis and
PT Ehrlichia canis
XX
PS Disclosure; Fig 24A; 55pp; English.
XX
CC The invention provides isolated outer membrane proteins (OMP) from
CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part
CC of the OMP family and consist of proteins OMP-1, -(B to Z) shown
CC in AAY06943-958. The E. canis proteins form part of the p30 family and
CC consist of proteins shown in AAY06959-970. The proteins and genes are
CC used to detect E. chaffeensis in patients and E. canis in dogs.
XX
SQ Sequence 831 BP; 300 A; 126 C; 136 G; 269 T; 0 other;

Alignment Scores:
Pred. No.: 5.82e-29 Length: 831
Score: 341.50 Matches: 90
Percent Similarity: 50.70% Conservative: 54
Best Local Similarity: 31.69% Mismatches: 117
Query Match: 22.62% Indels: 23
DB: 20 Gaps: 10

US-09-846-808-1 (1-284) x AAX34764 (1-831)
QY 7 ArgLysPheValLeuTrpValMetLeuIleLeuPheThr-----ProHisIleSer 23
DB 10 AAAAAATTTCTTATACAACTACATTTGGTATCTACTAACAACTCTTTTACCTGGCATATCT 69
QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
DB 70 TTCTCCAAACCAATACATGAAACAACTACTACAGGAACCTTTTACATATTATGGAAATAT 129
QY 41 LysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSerAlaAlaAsnThrVal 60
DB 130 GTACCAAGTATTTTCACATTTTGGGAACCTTTTACAGCTAAAGAACAAAAACAACACTACT 189
QY 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
DB 190 GGAATTTTGGATTAAGAAATCATCATGCTGGTGTATCATCTCTGAT---AAGAACAT 246
QY 81 ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
DB 247 GCAGCTTTTAAATATCCCAATATTATTCATTAATATGAAATATTAATCCATTTTAGGATTT 306
QY 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
DB 307 GCAGGGTAATTTGGCTATTCAATAGGTAGTCCA---AGAATACAATTTGAAGTATCATAC 363
QY 119 GluThrPheHisIleLysAsnGlyTyrLys---ArgIleAspCysGluLysHisPhe 137
DB 364 GAGACATTCGATGTACAAATCCAGAGATAAGTTTAAACAATGATGCACATAAGTATGT 423
QY 138 AlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnLysTyrValThr 157
DB 424 GCTTATCCATGAT-----TCCAGTAAACAACTGAAAGCTGTAATTCGGTTTT 474
QY 158 LeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAsp 177
DB 475 CTCAAAAATGAAGGATTAAGTGACATATCACATCATCTGTTAAATCTATGTTATGATAATA 534
QY 178 GlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAsp 197
DB 535 AACAAAAAGAAATGCTTTTTCACCTTACATATGTGCAGGCACTTGGTACTGACTTAATATTC 594
QY 198 PheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThr 217
DB 595 ATGTTTGACGCTATAAACCATCAAGCTGTTTATCAAGGAAATTAGGTTTATTAATCA 654
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 19, 2002, 03:00:25 ; Search time 1872 Seconds
(without alignments)
3174.752 Million cell updates/sec

Title: US-09-846-808-1
Perfect score: 1510
Sequence: 1 MSRRSRKRVLWMLILFTP.....ALAMLSIGYGGSIGIKFIL 284

Scoring table:
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n_model -DEV=xlh
-O=/cgn2.1/USPTO_spool/US09846808/runat_18102002_140435_26016/app_query.fasta_1.455
-DB=GenEmbl -QWRT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09846808 -CGN 1.1.1182 -runat_18102002_140435_26016 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: gb_wa:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1510	100.0	14759	1	AF230642 Ehrlichia
2	1510	100.0	27190	1	ECU72291 Ehrlichia c
3	932	61.7	28254	1	AF078553 Ehrlichia
4	448	29.7	11329	1	AF082744 Ehrlichia
5	425	28.1	6913	1	AF324792 Ehrlichia
6	424	28.1	3535	1	AF125274 Cowdria r
7	424	28.1	3538	1	AF125277 Cowdria r
8	424	28.1	3541	1	AF125279 Cowdria r
9	424	28.1	3572	1	AF125278 Cowdria r
10	422	27.9	3507	1	AF125276 Cowdria r
11	422	27.9	3551	1	AF125275 Cowdria r
12	411.5	27.3	843	6	AX042314 Sequence
13	411.5	27.3	4683	1	AF062761 Ehrlichia
14	397.5	26.3	1263	1	CRU50830 rum
15	393	26.0	1101	1	CRU50834 rum
16	392	26.0	1278	1	CRU50832 rum
17	391	25.9	1564	1	CRU49843 rum
18	389	25.8	873	1	AF368001 Cowdria r
19	389	25.8	828	1	AF368007 Cowdria r
20	388	25.8	828	1	AF368014 Cowdria r
21	388	25.7	1278	1	CRU50835 rum
22	387.5	25.7	1467	1	CRMAP1
23	387.5	25.6	1283	6	AX042313 Sequence
24	386	25.6	1307	1	AF077732 Ehrlichia
25	386	25.6	1309	1	AF393394 Ehrlichia
26	386	25.6	1309	1	AF393390 Ehrlichia
27	385	25.5	1243	1	AF077733 Ehrlichia
28	384	25.4	807	1	AF368004 Cowdria r
29	384	25.4	2977	1	AF068234 Ehrlichia
30	384	25.4	27190	1	ECU72291 Ehrlichia c
31	383.5	25.4	864	6	AX042305 Sequence
32	383.5	25.4	1265	1	CRU50833 rum
33	383.5	25.4	28254	1	AF078553 Ehrlichia
34	382	25.3	831	1	AF368000 Cowdria r
35	381.5	25.3	816	1	AF325176 Cowdria s
36	381.5	25.3	816	1	AF368013 Cowdria r
37	380	25.2	861	6	AX042312 Sequence
38	379.5	25.1	834	1	AF368011 Cowdria r
39	378.5	25.1	825	1	AF368008 Cowdria r
40	378.5	25.1	834	1	AF368010 Cowdria r
41	378	25.0	863	1	AF355200 Cowdria r
42	378	25.0	873	1	AY028378 Cowdria r
43	376.5	24.9	1342	1	AF393388 Ehrlichia
44	376.5	24.9	2362	1	AF319940 Cowdria r
45	376.5	24.9	11329	1	AF082744 Ehrlichia

ALIGNMENTS

RESULT 1
AF230642 AF230642 14759 bp DNA linear BCT 01-JUN-2000
LOCUS Ehrlichia chaffeensis clpx (clpx) gene, complete cds; 28-kda outer
DEFINITION Membrane protein gene cluster, partial sequence and unknown genes.
ACCESSION AF230642
VERSION AF230642.1 GI:8132823
KEYWORDS Ehrlichia chaffeensis.

ORGANISM	Ehrlichia chaffeensis	
REFERENCE	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;	
AUTHORS	Rickettsiaceae; Ehrlichiae; Ehrlichia; canis group.	
TITLE	Yu,X.J., McBride,J.W., Zhang,X. and Walker,D.H.	
JOURNAL	Characterization of the complete transcriptionally active Ehrlichia	
REFERENCE	Chaffeensis 28 kda outer membrane protein multigene family	
AUTHORS	Gene 248 (1-2), 29-68 (2000)	gene
TITLE	Yu,X.J., McBride,J.W., Zhang,X. and Walker,D.H.	
JOURNAL	Direct Submission	
FEATURES	Submitted (01-FEB-2000) Pathology, University of Texas Medical	
source	Branch, 301 Univ. Blvd., Galveston, TX 77555-0609, USA	
	Location/Qualifiers	
	1. .14759	
	/organism="Ehrlichia chaffeensis"	
	/strain="Arkansas"	
gene	/db_xref="taxon:945"	
	400. .1731	
CDS	/gene="clpX"	
	400. .1731	
	/gene="clpX"	
	/note="ATP-dependent protease"	
	/codon_start=1	
	/transl_table=11	
	/product="ClpX"	
	/protein_id="AAF73407.1"	
	/db_xref="GI:8132824"	
	/translation="MADNEKSCSFCGKIHSEVRKLIAGPSVFTCNCDLCSGIL	
	QESRSYKTDILKLPKEIKKVLDEYVIGQBSKKVLSVAVNYHKRLNSLVSIEV	
	EISKNVLDIGTGGTKLLARLQVLPFAMADATTLTEAGYVGEDVENILLKLL	
	QAANFNVDAAQGIYIDEVDIRKSENTSRDVSGEVQQAALKVIEGTVSVP	
	QGRHPHGSASLQVLMQLNSRKVSSTTKVTALYTIETIIQDFELSLQREV	
	LELNEGATRDVDQKQYEDNRYFALCKDSESTNSNGHVMVKNEGFRVSL	
	TNACYDMLVNSSLVPSACIGIGGITNFLAGTNIHTFLKALGIGFLISPKTVIFA	
	NGYVYKTNSTNLSVQVPLELKEAPKHIDPIACPNADNYGEGVGLRFL"	
	2236. .3051	
gene	/gene="p28-1"	
CDS	2236. .3051	
	/gene="p28-1"	
	/note="component of 28-kda outer membrane protein	
	multigene family"	
	/codon_start=1	
	/transl_table=11	
	/product="p28-1"	
	/protein_id="AAF73408.1"	
	/db_xref="GI:8132825"	
	/translation="MLELTPHISLASVLNDHNSMVVGIOYKPAROHLKLLIKESAA	
	NVVEFGLKDLNLLTGIKDNTNENIKYNNRNLGFGSIFGYNNKPNRIESE	
	LSYEFPHKNNGYKRIDCKEHFALAKEISGGSNNPANNKYVTLINNGISLTALINVC	
	YVDGLKHNIITYSCIGFGVDITDFLSKYTTTFYSQGLGASYVSPQSVFIEGYIH	
	GLFGKFKERIPINYPCDVPSPTPPNSKPHVHTALAMLGIGYVGSIGIKFIL"	
	3158. .3604	
CDS	/note="component of 28-kda outer membrane protein	
	multigene family"	
	/codon_start=1	
	/transl_table=11	
	/product="unknown"	
	/protein_id="AAF73409.1"	
	/db_xref="GI:8132826"	
	/translation="MNNKSRGOVLLFAAYSTFKYINCSCTSKRFSVKCFVYNPLKYF	
	YKSLFKLGKGFVNYNYSKTSNLRSLFKLFGIIRTNFKISYINFTNPRVLIIY	
	YKILKNNSINYLNNKTYNTATSPINSVIIGIYPGFVNNY"	
	3614. .4507	
gene	/gene="p28-2"	
CDS	3614. .4507	
	/gene="p28-2"	
	/note="component of 28-kda outer membrane protein	
	multigene family"	
	/codon_start=1	
	/transl_table=11	
	/product="p28-2"	
	/protein_id="AAF73410.1"	
	/translation="MSYAKVFIILICILILLVPSLSFAIVNNDFLKDNIGHIFYIGQYKP	
	GYPRFNRLVTNNIRELMSSDECRSTIPHMVQSVAGQGLPPEALEELADGFPQY	
	LYETIDYNTYKKNLILGAGVIGYSTHERVEVEAFYDKFENLTAPAGYLHKNFYEFA	
	LATYMTKPHQSAEDKYYMYKNTGITLSPFLINACYDFILKTRNAPYVLCILGVGN	
	FDLQVFSKFAIQAQVGLSYFVSPNIAIFFIDGSPHGLNNQFSDSPVVDYSSGFP	
	TISAKFNANFLTSSIGIRFIS"	
	4524. .5381	
gene	/gene="p28-3"	
CDS	4524. .5381	
	/gene="p28-3"	
	/note="component of 28-kda outer membrane protein	
	multigene family"	
	/codon_start=1	
	/transl_table=11	
	/product="p28-3"	
	/protein_id="AAF73411.1"	
	/db_xref="GI:8132828"	
	/translation="MQKLIISFILSGLLLPKYVFCMHQNNIDGSYVTIKYQLTTPH	
	FKNFYIKETDFTQEPDIGLAKITANTKFDTLKENFSFSLHOTDSYKSYQNDLLGIGL	
	SVGLFVKFRIEPEGAYKNFNTKRLARYKSKDGKYFAIPRKSEHGFLDNKFGYTVAK	
	NGGISLIISNIIINLCSETKYKSFYICIGVGGDFIEFDVMRIKFAOVGVSYPT	
	SKLILSINGOYHKVIGNKPELLPYQPVVELKRLVTKTSKDIDQDVTASITLNEHFS	
	SEIGLSFIF"	
	5394. .6212	
gene	/gene="p28-4"	
CDS	5394. .6212	
	/gene="p28-4"	
	/note="component of 28-kda outer membrane protein	
	multigene family"	
	/codon_start=1	
	/transl_table=11	
	/product="p28-4"	
	/protein_id="AAF73412.1"	
	/db_xref="GI:8132829"	
	/translation="MYMYKKHYCIYTVYITLFFLLPIESLSALIGNVEKDLKVSS	
	TVVSSQKPSIFHRNFSIQESHPKSSSEEFKIKANLNILKSNAYNLQFQNTTST	
	SETGYFSGGLRLEAGCYQEFNVKNSNLSLISSNKYHSRIHDEYAITNNKLSFA	
	SIMVNTCYDISINNTSIVPYLCGTIGEDLVGLFNTIHFKLAYQGVGMSYLINNILL	
	FSDIYYHKVMGNRFKNLYMQYVADPNISEITIPILAKLDIGYFGSGEIGRFEN"	
	6228. .7115	
gene	/gene="p28-5"	
CDS	6228. .7115	
	/gene="p28-5"	
	/note="component of 28-kda outer membrane protein	
	multigene family"	
	/codon_start=1	
	/transl_table=11	
	/product="p28-5"	
	/protein_id="AAF73413.1"	
	/db_xref="GI:8132830"	
	/translation="MTKKFNVMVILTLFLFLPKLSFTTYANNNTITQKVGLYISGQ	
	YKPSIPHFKNFSVEENDKVVDLIGLTDVTYITEHILRDNTKENTHYIAKFNKINF	
	SSAIGYYSQGPRLIEISSYGDYDFVYKNYAVQDYNRYFALYREKNGSNFSPKHET	
	SQPSDSNPKRSFYLLMKNGVFAVSIINGCYDFSPNNTTISPYICIGVGGDFIEFE	
	VMHIFACQSKVGISYIPIPSITIFADAHYHKVINNNKFNHNLHVKYSYELKNSPTTISA	
	TAKLNIEYFGGEVGMRFIF"	
	7137. .7976	
gene	/gene="p28-6"	
CDS	7137. .7976	
	/gene="p28-6"	
	/note="component of 28-kda outer membrane protein	
	multigene family"	
	/codon_start=1	
	/transl_table=11	
	/product="p28-6"	
	/protein_id="AAF73414.1"	
	/db_xref="GI:8132831"	
	/translation="MSKKPFTITGTVLASLLSFLSIESFSAINHNHTGNNTSGIYITG	
	OYRPGVSHFSNFSVKNVDYIOLGVKKKSASSIDPNTYSNFOGYPYTVTFODNAASF	
	GAGYSYPSLSRLESGSEKFDKPKDYSAKDAFFFPALARTSTTVPODAOKYTVM	
	KNNGLSVASIMINGCDLSENNLVSPYICAGIGEDFIEFFDPLHLKLAYQGLGISY	
	YFPKINVFAGGYHVRIGNKFNLVNHHVVTDPDEFPKATSAVATLNVATFVGEGAVK	

CDS	8214.. .8591	FTF"	
	/note="Orf378"		
	/codon_start=1		
	/transl_table=11		
	/product="unknown"		
	/protein_id="AAF73415.1"		
	/db_xref="GI:8132832"		
gene	/translation="MNCYLKGLPLNLOIFYGIFSRNFQNNLTLLIIPNDKSGFYTTL		
	WDPALHYTVTLGSEVSENFEDILYENIQCKLLINYSVLNQHKNKTLVLIPIPN		
	ARFSEIRVRNYSINKESSEC"		
	8581.. .9432		
	/gene="p28-7"		
CDS	8581.. .9432		
	/gene="p28-7"		
	/note="component of 28-kDa outer membrane protein		
	multigene family"		
	/codon_start=1		
	/transl_table=11		
	/product="p28-7"		
	/protein_id="AAF73416.1"		
	/db_xref="GI:8132833"		
	/translation="MSAKKKLFIIIGSVLVCLVSLPDKSLNLSNNINNNKCTGLYVS		
	QKAPTWSHFSNLSKETYTDTELLGLADIKISDITITNKKFNIPYNTKFDQNAVS		
	FSAAGVYSQDSRVEVEWSEYEDFVNKPNYVYSEAFRYIALARGIDNLOKQYPTNK		
	YVINKNLSVASIINGCYDFSLNNLKVSPIYICVGFEGDIIIEFFSAVSFKFAYQGRV		
	GISVPLFSNMIIFADGYVHKVIGNKENNLNVQHVVSLSNHPKSTFAVATLNVEYFGE		
	FGLKFI"		
	9441.. .10268		
gene	/gene="p28-8"		
CDS	9441.. .10268		
	/gene="p28-8"		
	/note="component of 28-kDa outer membrane protein		
	multigene family"		
	/codon_start=1		
	/transl_table=11		
	/product="p28-8"		
	/protein_id="AAF73417.1"		
	/db_xref="GI:8132834"		
	/translation="MSKKNFTTIGATLIHMLPNISFPETINNNTDKLSGLYISGOYK		
	PGTSHFSKSVKETYNDNIOLIGLRHNAISTSLNINTDFNIPVKVYFQNNITFSFGA		
	IGSDPTGARFEELGSEEDVDGDCCLIKDYRYEALARNSGSSPTSNNTVMRN		
	DGVSTVIFNGCDDIFLKDLEVSYPYCVGVGGDFIEFFDHALIKLAYQKGLKGINYHL		
	STQASVFIDGYHKVIGNQPNLNVQHVASTDFGPVYAVATLNIYFGGEIGIRLTF"		
	10278.. .11135		
gene	/gene="p28-9"		
CDS	10278.. .11135		
Alignment Scores:			
Pred. No.:	2,6e-126	Length:	14759
Score:	1510.00	Matches:	284
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0
US-09-846-808-1 (1-284) x AF230642 (1-14759)			
Qy	1	MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuIleLeuPheThrPro	20
Db	2197	ATGAGTAAGAGGATACAGAAAGTTTCTTTGGGGTATGTTAAATATTATTACACCA	2256
Qy	21	HisIleSerLeuAlaSerValLeuAsnAspHisAsnSerMetTyrValcIlyleGlnTyr	40
Db	2257	CATATTTCTTAGCATCAGTTTAAATGACCATTAATCTATGATGTTGGTATTCAGTAC	2316
Qy	41	LysProAlaArgGlnHisLeuSerLysLeuIleLeuLysGloSerAlaAlaAsnThrVal	60
Db	2317	AAGCCAGCCAGACACACTCTCAAAACTTCTTATAAAGAAAGCGCAGCCAACTGTA	2376
Qy	61	GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuThrGlyIleLysAspAsn	80
Db	2377	GAAGTTTGGGTTAAAAAGAGACTTACTAAATGATCTATTAACTGGTATTAAAGATAAT	2436

Qy	81	ThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeuGlyPheSerGly	100
Db	2437	ACTAATTTTAAATATAAATACAAACCATATTATGAAATAATATCGTTTAGTGTTCCTCGGT	2496
Qy	101	IlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThr	120
Db	2497	ATATTGGATATTACTATAACAAAAATTTTCAGAAATAGAACTCTGAACCTTTCTTTACGAAACT	2556
Qy	121	PheHisIleLysAsnAsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAla	140
Db	2557	TTCCACATAAAGAACACGATATAAAGAATTTGATTTGTGAAGACATTTTCTCTCTGCA	2616
Qy	141	LysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsn	160
Db	2617	AAGAAATAAGCGGTGGTAGTAACAATCTCGCAATAATAATAATATGTTACCCCTAATAAAT	2676
Qy	161	AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLys	180
Db	2677	AATGGTATTTCACCTTACCTCAGCTCGATCAACGTTTGTATGATGTTGACGACTTAAA	2736
Qy	181	HisAsnIleLeuThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSer	200
Db	2737	CATAATAATAACGATATTTCATGTTTAGGATTTGGAGTAGACACAATTTGATTTTCTCAGT	2796
Qy	201	LysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerPro	220
Db	2797	AAGTATACTACAAAGTTTTCATACCAAGAAATATGAGGAGCTAGTTATATCTGTTCTCCT	2856
Qy	221	GlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGlyLys	240
Db	2857	CAAGTATCGGTTTTTATAGAAGGTTATTATCATGTTTTATTGGTAAAAAATTTGAGAAA	2916
Qy	241	IleProValAsnTyrProCysAspTyrProSerProThrProAsnSerLysProHis	260
Db	2917	ATTCTCTGTAATTTATCCATGTCATTCATCATCCACCAACACCAACAGCAGCCACAT	2976
Qy	261	ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlycylSerIleGlyIle	280
Db	2977	GTTACACTACAGCATTTAGCTATGTTAACTATTGGATATTACGGTGGGAGTATTGAATA	3036
Qy	281	LysPheIleLeu	284
Db	3037	AAGTTTATATTA	3048
RESULT 2			
LOCUS	ECU72291	27190 bp	DNA linear BCT 02-APR-2001
DEFINITION	Ehrlichia chaffeensis strain Arkansas major outer membrane protein		
ACCESSION	U72291 AF021338		
VERSION	U72291.2		
KEYWORDS	GI:13511827		
SOURCE	Ehrlichia chaffeensis.		
ORGANISM	Ehrlichia chaffeensis.		
REFERENCE	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales; Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.		
AUTHORS	1 (bases 14844 to 21136; 21479 to 22334)		
TITLE	Immunodominant major outer membrane proteins of Ehrlichia chaffeensis are encoded by a polymorphic multigene family		
JOURNAL	Infect. Immun. 66 (1), 132-139 (1998)		
MEDLINE	98084465		
PUBMED	9423849		
REFERENCE	2 (bases 1 to 27190)		
AUTHORS	Ohashi, N., Rikihisa, Y. and Unver, A.		
TITLE	Analysis of transcriptionally active gene clusters of major outer membrane protein multigene family in Ehrlichia canis and E. chaffeensis		
JOURNAL	Infect. Immun. 69 (4), 2083-2091 (2001)		
MEDLINE	21153566		
PUBMED	11254561		
REFERENCE	3 (bases 1 to 27190)		

AUTHORS TITLE JOURNAL	Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y. Direct Submission Submitted (24-SEP-1996) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA	gene	/translation="MNNKSRGQVLLFAAYSIKFYINSCTSKFRLSVKCFVYNPLKIFYKTLKLGKGFVNNYKSTSNLHRSFLKFLGIIRTNFKISYLNFTNPRVLIIYKILKKNKNSSINYLNNKYTINPTATSFINSVLIIGIYPCFVNNY"
REFERENCE AUTHORS TITLE JOURNAL	4 (bases 1 to 27190) Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y. Direct Submission Submitted (27-AUG-1997) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA	CDS	/gene="omp-10" 3915..4808 /note="omp-1 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein OMP-1Q" /protein_id="AAK28661.1" /db_xref="GI:13511832"
REFERENCE AUTHORS TITLE JOURNAL	5 (bases 1 to 27190) Ohashi,N., Rikihisa,Y. and Unver,A. Direct Submission Submitted (29-NOV-2000) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA	gene	/translation="MSYAKVFILICILLVPSLSFAIVNDFLKDNIHGFIYIGQYKPGVPRNRFVLTNNNIRELSSDEECSTIPHMVQSAQGTLPPEALEELAKGLHGYYLFTLPNTYKKNLLGAGVIGYSTHERVEVEAEYERFNLTAAGYLHKNFYEFALATMTDKHPQSAEDKYIYMKNTGITLSPFIINACYDFILKTRNVAPLYCLGVGGNFIDLDQVSKFAYQAKVGISYFVSPNIATFFIDGSPFHGLNNOFSDLPVVYSSSGFPTISAKFNANFLTSSIGIRFIS"
REMARK COMMENT	Sequence update by submitter On Apr 2, 2001 this sequence version replaced gi:2853584 gi:2853273.	CDS	/gene="omp-1p" 4825..5682 /note="omp-1 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein OMP-1p" /protein_id="AAK28662.1" /db_xref="GI:13511833"
FEATURES Source	Location/Qualifiers 1..27190 /organism="Ehrlichia chaffeensis" /strain="Arkansas" /db_xref="taxon:945" 315..956 /codon_start=1 /transl_table=11 /product="hypothetical transcriptional regulator"	gene	/translation="MOKLYISFILSGLLPKYVFCMQHNNNIDSYVTIKYQLTTPHFKNFYKEDFTDQEPDGLAKITANTKFTDLKENFSPLHOTDSKSYQNDLLGIGLKNFYKFKRIEFDGAYKNFTKRLARYKSKDGYKYPAPRKSEHGFLODNTFCYTVAKNNGIISINILNLCSTKYKATPYICIGVGDFIETFDVMBRIKFAQCKVGYSPITSKLISINQYHKVIGNKFELLPVQPVVELKRLVTKT7SKDIOQDVTASLTMLEHFSSEIGUSFIF"
gene	1136..2032	CDS	/gene="omp-1t" 5695..6513 /note="omp-1 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein OMP-1t" /protein_id="AAK28663.1" /db_xref="GI:13511834"
gene	2498..3352	gene	/translation="MYMKKKHYCYIVTVITLFLFLLPDLIESLSALIGNVEKDLKYSSTVSSQYKPSIFHFNFISOEHPKSSSEFKKANKANLNILKSAYINQFQDNTTSFSGTIGYFSGLEAEAGCYOEFNVKNSNLSIISSNKYHSRIHDEYAITNNKLISIASIMVNTFYDISINNTSIVPYLCTGIGEDLVGLFNTTHFKLAYQGVKGYMLINNILLFSDIYYHKYMGNRKKNLYMQYVADPNISEETIPIAKLDIGYFGSGEIGIRFMEN"
gene	2498..3352	CDS	/gene="omp-1u" 6529..7416 /note="omp-1 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein OMP-1u" /protein_id="AAK28664.1" /db_xref="GI:13511835"
gene	3459..3905	gene	/translation="MTKFNFFNVNLTLLFLPLKSFYYANNNTITQKVGLYISGQYKPSIPHFKNFSVEENDKVDLIGLTDVITYTEHTLRONTKFNTHYIAKFKNNFIETSSAIGYVSGQGRLEIETESSYGDVNVKNYAQDVNRYFALVREKNGSNFSPKPHETSPQSDNPKSKSYTLLMKNNGYFVASVIINGCDYFSFNNTTISPIYICIGVGGDIERFEVWHIRFACQSKVGISYIPISPSITIFADAHYHKVNNKFNLNHVKIYSTELKNSPTITISA TAKLNIEYFGGEGVMRFIF"
CDS	7438..8277	CDS	/gene="omp-1v" 7438..8277 /note="omp-1 family member" /codon_start=1 /transl_table=11

/product="major outer membrane protein omp-1w"
/protein_id="AAK28665.1"
/db_xref="GI:13511836"

/translation="MSKKKEITIGTVLASLLSFLSIESFSAINHNHTGNNTPGIYITG
QYRGVSHFSNFKETNVDIQLGVGKKSSIDPNYFNQFGPIYTFODNRAASF
GAIGTSPESURLESGYERKFDVDPKDYSAKDAFFALARNITSTVPDAQKITVM
KNGLSVASIMINGCYDLSFNVLVSPYICAGIGEDFEFFDTLHLKAYQCKLGISY
YFFPKINVAGGYHVRVIGNFKNLNVNHVVTLDDEPKATSAVATLNVAYFGGAGVK
FTF"

CDS 8300..8890

/codon_start=1

/transl_table=11

/product="unknown function UN3"

/protein_id="AAK28666.1"

/db_xref="GI:13511837"

/translation="MIYKEKLTTRVGEYLAVLSFTLSTYIFLVLNIIIRNSLAICVI
SLLTNIFNVSTKLLDKCRDTKFSNMNXYLKGPLNQLQIFYGFIRNFQNTLI
IPNSKCGFYTLMDNPNALHYTLTSEYRNFFDILYENIICOCKLLINRNSVLNQ
HNKNTLVIIPIPNAREFSNEIRVRNISINKESSEYEC"
8880..9731

gene /gene="omp-1w"

CDS 8880..9731

/gene="omp-1w"

/note="omp-1 family member"

/codon_start=1

/transl_table=11

/product="major outer membrane protein omp-1w"

/protein_id="AAK28667.1"

Alignment Scores:

Pred. No.:	5,45e-126	Length:	27190
Score:	1510.00	Matches:	284
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-846-808-1 (1-284) x ECU72291 (1-27190)

Qy	1	MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuLeuPheThrPro	20
Db	2498	ATGAGTAAGAAGAGTAAACAGAAGTTGTTTGTGGGTATGTTAATATTATTTACACCA	2557
Qy	21	HisIleSerLeuAlaSerValLeuAsnAspHisAsnSerMetTyrValcIlyleGlnTyr	40
Db	2558	CATATTTCTTTAGCATCAGTTTAAATGACCATAATCTATCTATGTTGGTATTCAGTAC	2617
Qy	41	LysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSerAlaAlaSerThrVal	60
Db	2618	AAGCCAGCCAGACACCTATCAAACTCTTATATAAAGAAAGCCAGCCACACTGTA	2677
Qy	61	GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyTleLysAspAsn	80
Db	2678	GAAGTTTGGGTTAAAAAGACCTACTAATATGATCTATTAACTGGTATTAAAGATAAT	2737
Qy	81	ThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeuGlyPheSerGly	100
Db	2738	ACTAATTTTAATATAAATACACCATATTATGAAATAATACGTTTAGGTTTTCCGGT	2797
Qy	101	IlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThr	120
Db	2798	ATATTGGGATATTACTATAACAAAAAATTCAGAATAGAACTCGAACTTCTTACGAAACT	2857
Qy	121	PheHisIleLysAsnAsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAla	140
Db	2858	TTCCACATAAAGAACACGGATATAAAGAAATGATGTGGAAGCAATTTGCTCTTGA	2917
Qy	141	LysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsn	160
Db	2918	AAGGAAATAAGCGGTGGTAGTAACAATCCTGCCAATAATAATATGTTACCTTAATAAT	2977
Qy	161	AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLys	180

Db	2978	AATGGTATTTCACCTACCTCAGCTCTGATCAACGGTTTGTATGATGTTGACGGACTTAA	3037
Qy	181	HisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSer	200
Db	3038	CATAATATAAATACGTATTCACTGTTTAGGATTTGGAGTAGACACAAATGATTTTCAGT	3097
Qy	201	LysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyValAspSerTyrThrValSerPro	220
Db	3098	AAGTATACTACAAAGTTTTCATACCAAGCAAAATAGGAGCTAGTTATATACTGTTCTCTCT	3157
Qy	221	GlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLys	240
Db	3158	CAAGTATCGGTTTTATAGAAGCTTATTATCATCTGTTTATTTGGTAAAAAATTTGAGAA	3217
Qy	241	IleProValAsnTyrProCysAspTyrProSerProThrProProAsnSerLysProHis	260
Db	3218	ATTCTGTAAATATCCATGCTGACTATCCATCACCACCAACCAACAGCAAGCCACAT	3277
Qy	261	ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyIle	280
Db	3278	GTTCACTACAGCATTAGCTATGTTAAGTATTGGATATTACGGTGGAGTATTGGAATA	3337
Qy	281	LysPheIleLeu	284
Db	3338	AAGTTTATATTA	3349
RESULT	3		
LOCUS	AF078553	28254 bp	DNA linear BCT 02-APR-2001
DEFINITION	Ehrlichia canis major outer membrane protein p30 multigene cluster 1, complete sequence.		
ACCESSION	AF078553 AF078554 AF078555 AH006958		
VERSION	AF078553.2 GI:13512584		
KEYWORDS	Ehrlichia canis.		
SOURCE	Ehrlichia canis.		
ORGANISM	Ehrlichia canis.		
REFERENCE	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales; Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.		
AUTHORS	1 (bases 1 to 28254)		
TITLE	Ohashi,N., Unver,A., Zhi,N. and Rikihisa,Y. Cloning and characterization of multigenes encoding the immunodominant 30-kilodalton major outer membrane proteins of Ehrlichia canis and application of the recombinant protein for serodiagnosis		
JOURNAL	J. Clin. Microbiol. 36 (9), 2671-2680 (1998)		
MEDLINE	98371112		
PUBMED	9705412		
REFERENCE	2 (bases 1 to 28254)		
AUTHORS	Ohashi,N., Rikihisa,Y. and Unver,A.		
TITLE	Analysis of Transcriptionally Active Gene Clusters of Major Outer Membrane Protein Multigene Family in Ehrlichia canis and E. Chaffeensis		
JOURNAL	Infect. Immun. 69 (4), 2083-2091 (2001)		
MEDLINE	21153566		
PUBMED	11254561		
REFERENCE	3 (bases 1 to 28254)		
AUTHORS	Ohashi,N., Unver,A., Zhi,N. and Rikihisa,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JUL-1998) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Road, Columbus, OH 43210, USA		
REFERENCE	4 (bases 1 to 28254)		
AUTHORS	Ohashi,N., Rikihisa,Y. and Unver,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-NOV-2000) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Road, Columbus, OH 43210, USA		
REMARK	Sequence update by submitter		
COMMENT	On Apr 2, 2001 this sequence version replaced gi:3790556 gi:3790555		
FEATURES	Location/Qualifiers		
source	1..28254		
	/organism="Ehrlichia canis"		
	/strain="Oklahoma"		
	/db_xref="taxon:944"		

CDS	<1. .615 /codon_start=1 /transl_table=11 /product="hypothetical transcriptional regulator" /protein_id="AAK28679.1" /db_xref="GI:13512585" /translation="PYANEIKAHDPDIEDIEKNIOQHKTGEGRLDVEDYASKNF KEGASTRELKATVRPHDPVDECYGEKIKRORIMRGMSONLANKLGITFOOVOKY EKGTRIVISRLYQLASVLNVEVRDMLKLQEDLKNISCDNPITPPHALNDNEKEFLP EPNDSIDSKVELMMVRATCTICKERKVRNIYINLVKALSLDNKS" 843. .1733 /gene="p30-19" /gene="p30-19" /note="p30 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein p30-19" /protein_id="AAK28680.1" /db_xref="GI:13512586" /translation="MKLLYHLDNIMIKSAIGIVFVSFIALFAPNAPSPVPIDFSNES EMAGFYAQYINIGFRPFSISAKYTKDEKSEKELTFLSKEETETIDUKKAGDFRKG YSYVARNYTGSGAIGYGGGLRVLEGSFTFRDYDKQKKNPDGHRYPALSKDSEI QNSGSSGNNKDYVVMKNEGFAISLMFNACVDYMIIGNSSLVPNACIGIGQGIIRFL QGTNIHTLAKLGLGLISPKTILFANGYVYKAKDNAPTNLVQYVPEISAAPKHID PIVYNADNVCCEVGLRFL" 2491. .3318 /gene="p30-18" 2491. .3318 /gene="p30-18" /note="p30 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein p30-18" /protein_id="AAK28681.1" /db_xref="GI:13512587" /translation="MNNKLSLYIALILFTSHVSSALVNDHNLVYFGYKPKARHHL SNLLIKESQDVVEVLAKYDAIGSPDSTKEVNNFTIKYPHVDNNRLGFSVIFGYV YKNFRIESEIHEIFQLNEGHRVGPKEYPALKAPPSSTQGYRHYVLINNGISTY SALINACYDLVIPAHHIITYSCIGFDIVIDFLSKYTRFSGQKLGASPTPIHRMSV FTEVYVHGFLGKFEQLPLNLYNANTSPPOQPHVHTTASAILSIGYVGSVGKIFIL" 3425. .3871 /codon_start=1 /transl_table=11 /product="unknown function U2" /protein_id="AAK28682.1" /db_xref="GI:13512588" /translation="MNNKVLQNEVLVLLALPYSLHNSIDIEYNEQPIKKLSRLNKNTI KSQKLHNLKYILLKLKTISSIGWFLKFSNDYSNQVIYVYLRTSLSKLVSITIALDFDK NIFLNYKVLTDLSLRNFSANYFYSKHHNITNCFVSRAYVFLSCSY" 3882. .4832 /gene="p30-17" 3882. .4832 /gene="p30-17" /note="p30 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein p30-17" /protein_id="AAK28683.1" /db_xref="GI:13512589" /translation="MNYAKVFILVCIIFLPPSLFSFATNNNYPVHEIGKSGHGFYGVQ YKPGTPHFRFSIADSTFNMLIAISHTKDYLFYSYSTEVRKGLFSLPQEQNLLHYATGG STLNTLKSNKFIPOGNYPTDNLGVGIVGYSINNLRIELEAFYKFNKAPGKY NYDTEYATLVVYKGTQPVHYCHMKNTGIILSSVLVATCYDFTLKIAPYLCL VGVGDIFELQTRLRASQVAKGLYALSPNLTFFVDGSGFHGMNNOFPGLLVDPYPT DISVSGSDGNATYASEFTTMLAKLNMIFLAGSIGIRFS" 4848. .5699 /gene="p30-16" 4848. .5699 /gene="p30-16" /note="p30 family member" /codon_start=1 /transl_table=11	gene
CDS	5716. .6510 /gene="p30-15" 5716. .6510 /gene="p30-15" /note="p30 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein p30-15" /protein_id="AAK28685.1" /db_xref="GI:13512591" /translation="MVNLIINVIKYTIALAFLLLPRVSFSLICNIEKSIKLLSVHTN SQKPSISQISNYLIQENSKKKNILNLSNNTITYNMQLENSTNFRFIIGYFFKR LRFVADSYEEFHDKDNLSKANLSKYRMYNEDFONFTIADNKLUSITSAIVNICY DILINNTVPLHCTAVGICSTGFFNDMRKLLYQRIKGLYLINSNMVLEFFNYYHK VMRNKLNLLTQYSVGINAFDAITVLTANDDIGYFGEVGVRFIF" 6525. .7394 /gene="p30-14" 6525. .7394 /gene="p30-14" /note="p30 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein p30-14" /protein_id="AAK28686.1" /db_xref="GI:13512592" /translation="MLQRLNFIIILAFLLLLFPFQSFTLYIHDEHTQNVGLYISQ YKPSIDYFKNFLIEENSHKTVELMGANDVHTVTEYVKDNTKFTNTPYSKFRNSLIN LSGAIYSGOGPRLTEIEGYENFVASCNCPVKANRYIALVRDKKPCNIVPODHS HSNMSYFTFKNGCISTLSVMINGCYDIAESNVKISPYVCAGIGGDPITITFERMHIKF AYCGKGISILVSPSISIFANGHYHKVMDNVFNKHLVKYIYIKLQDAPTIISARAKLRI YCGKSEVGVRFVF" 7419. .8255 /gene="p30-13" 7419. .8255 /gene="p30-13" /note="p30 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein p30-13" /protein_id="AAK28687.1" /db_xref="GI:13512593" /translation="MNNKSLILGTLISLFSPIKAFSVINHSIDSSNVNGLYFTG OYRPAYSHESGFTVRETNIAQOLVSLNTKNENHIIERTNFSGLYAKFODNRAAPS GAIGSYPEGLKFEIEISYKFGVKSKNYQSTNAVIFALARQTTSSNPSDNKYVMHK NSGLSVASVMINGCYNMSFNLVSPYICAGIGEDFIEFFDTLYIKLAYOGKLGVNYS LSSRFNIFADMYHKVIGNQFKNLNVIHAVALDTPPKVTSAIATLNAYFGEVGIHF IL" 8266. .8823 /codon_start=1 /transl_table=11 /product="unknown function U3" /protein_id="AAK28688.1" /db_xref="GI:13512594" /translation="MATYTHKAKITPEMPSPSLTDYIILSLNNVODHNSFTICIT KLYKLULSSITKKLINVKDNIAVLHNLTYLNKCSKFSQTSQSEYEMFYTKNIKOR ILNFSLDIFIKYNVKNKSNESYKSYKSHNFNAYITTSFIHQHLLKYKYKYSTNOH NKSTKFSNKLAKNILLNNOENSYE" 8816. .9685 /gene="p30-12" 8816. .9685 /gene="p30-12" /note="p30 family member" /codon_start=1 /transl_table=11	gene
CDS	8266. .8823 /codon_start=1 /transl_table=11 /product="unknown function U3" /protein_id="AAK28688.1" /db_xref="GI:13512594" /translation="MATYTHKAKITPEMPSPSLTDYIILSLNNVODHNSFTICIT KLYKLULSSITKKLINVKDNIAVLHNLTYLNKCSKFSQTSQSEYEMFYTKNIKOR ILNFSLDIFIKYNVKNKSNESYKSYKSHNFNAYITTSFIHQHLLKYKYKYSTNOH NKSTKFSNKLAKNILLNNOENSYE" 8816. .9685 /gene="p30-12" 8816. .9685 /gene="p30-12" /note="p30 family member" /codon_start=1 /transl_table=11	gene

/product="major outer membrane protein p30-12"
/protein_id="AAK28689.1"
/db_xref="GI:13512595"

/translation="MNSKRTSILGSILICLAACLPQISFSESSNVNTYNTKHTGLYIS
GLYKRSVSHFSDFSKERYNTFEALFGKQDSSILRNKRTQYNNRNVPYTKAFQD
DFASPIAVGIANNPRIEESYEDVKNPGNYTIDAHRYTALAREKTSYLLS
PKENRYVTKNGISIVSIINGCYDILSNDSPKSPYICTGGDFIEFFSAIRKFA
YQKIGISYSLSNIIILTDGYYHKVINSQFNKLVNVEHVNELTDPKVTSAFLNI
EYFGEGFLKLF"

Alignment Scores:

Pred. No.: 1.08e-73 Length: 28254
Score: 932.00 Matches: 180
Percent Similarity: 77.14% Conservative: 36
Best Local Similarity: 64.29% Mismatches: 58
Query Match: 61.72% Indels: 6
DB: 1 Gaps: 2

US-09-846-808-1 (1-284) x AF078553 (1-28254)

Qy	5	SerAsnArgLysPheValLeuTrpValMetLeuLeuPheThrProHisIleSerLeu	24
Db	2494	ATAATAAACTTCCCTCTTATATGCTAATAATATATATACACATGTTCTTCA	2553
Qy	25	AlaSerValLeuAsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArg	44
Db	2554	GCATTGGTTTAAACGATCACAATCTGTATATTTGGTATTCAATATAAACGAGTAGG	2613
Qy	45	GlnHisLeuSerLysLeuLeuLysGluSerAlaAlaAsnThrValGluValPheGly	64
Db	2614	CATCATCTATCAATCTCTTATCAAGAAAGTAAGTCAGATGTTGTAGAAAGTCGTGCA	2673
Qy	65	LeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsn	84
Db	2674	CTGAATATGATGCATAGTAGTCCATGGACAGCAGTAAAGAGTAAATATTTTACC	2733
Qy	85	IleLysTyrAsnProTyrTyrGluAsnAsnArgLeuGlyPheSerGlyIlePheGlyTyr	104
Db	2734	ATAAAATACAAATCCACATATGATAATAATAGTTAGGTTCTCTGTAATATTTGCTAT	2793
Qy	105	TyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLys	124
Db	2794	TACTACAAATAAAATTTTAGGATAGAAATCCAGAAATTTCCACGAAATTTCCAGCTAAA	2853
Qy	125	AsnAsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSer	144
Db	2854	ATGAAGGACATAAAGAGCTTCGATTCGAGAAATATTTTGCACTAAG-----TTT	2904
Qy	145	GlyGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSer	164
Db	2905	GCTCCACCGCTCATCAGCAAGGATAGACATGTTACTTTAATAAACAATGATTTTCA	2964
Qy	165	LeuThrSerAlaLeuIleAsnValCysTyrAspValAspGlyLeuLysHisAsnIleIle	184
Db	2965	ACCACCTTCAGCTTAAATTAATGCTGCTATGATGACTCATACCTGCACATATAATAA	3024
Qy	185	ThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThr	204
Db	3025	ACATATTATGTTAGGGTTTGAATAGATAGATAGATTTTCTAAGTAAATACACTACA	3084
Qy	205	LysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerVal	224
Db	3085	AAATTTTCACATCAAGCAAGCTAGGAGCTAGTACCCTATTCTCATAGATGTCAGTC	3144
Qy	225	PheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsn	244
Db	3145	TTTACAGAGGTTTATACCATGTTTATTCGGTAAATAATTTGACGAGCTTCTCTTAAAC	3204
Qy	245	TyrProCysAspTyrProSerProThrProProAsnSerLysProHisValHisThrThr	264
Db	3205	TATAATGCT-----AATACATCCACCACCAACACACACGCTACGACT	3255
Qy	265	AlaLeuAlaMetLeuSerIleGlyTyrTyrGlyLysSerIleGlyIleLysPheIleLeu	284

Db	3256	GCATCAGCTATATTAAGTATTGGCTATTATGGTGAAGTGTTCGAATAAAGTTTATATTG	3315
RESULT	4		
AF082744			
LOCUS		11329 bp DNA linear BCT 18-SEP-2000	
DEFINITION		Ehrlichia canis p28 multigene locus, partial sequence.	
ACCESSION		AF082744 AF168788 AF168789	
VERSION		AF082744.2 GI:10181081	
KEYWORDS			
SOURCE		Ehrlichia canis.	
ORGANISM		Ehrlichia canis	
REFERENCE		Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;	
AUTHORS		Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.	
TITLE		1 (bases 1 to 11329)	
JOURNAL		McBride,J.W., Yu,X.J. and Walker,D.H.	
MEDLINE		28-kilodalton protein of Ehrlichia canis: a potential	
PUBMED		serodiagnostic antigen	
REFERENCE		Clin. Diagn. Lab. Immunol. 6 (3), 392-399 (1999)	
AUTHORS		99242757	
TITLE		10225842	
JOURNAL		2 (bases 1 to 11329)	
MEDLINE		McBride,J.W., Yu,X.J. and Walker,D.H.	
PUBMED		A conserved, transcriptionally active p28 multigene locus of	
REFERENCE		ehrlichia canis	
AUTHORS		Gene 254 (1-2), 245-252 (2000)	
TITLE		20432107	
JOURNAL		3 (bases 1 to 11329)	
MEDLINE		McBride,J.W., Yu,X.J. and Walker,D.H.	
PUBMED		Direct Submission	
REFERENCE		Submitted (07-AUG-1998) Pathology, University of Texas Medical	
AUTHORS		Branch, 301 University Blvd., Galveston, TX 77555-0609, USA	
TITLE		4 (bases 1 to 11329)	
JOURNAL		McBride,J.W., Yu,X.J. and Walker,D.H.	
MEDLINE		Direct Submission	
PUBMED		Submitted (04-AUG-2000) Pathology, University of Texas Medical	
REFERENCE		Branch, 301 University Blvd., Galveston, TX 77555-0609, USA	
AUTHORS		Sequence update by submitter	
TITLE		On Sep 18, 2000 this sequence version replaced gi:3769522.	
JOURNAL		Location/Qualifiers	
MEDLINE		1. .11329	
PUBMED		/organism="Ehrlichia canis"	
REFERENCE		/strain="Jake"	
AUTHORS		/db_xref="taxon:944"	
TITLE		<1. .146	
JOURNAL		/gene="p28a"	
MEDLINE		<1. .146	
PUBMED		/gene="p28a"	
REFERENCE		/note="outer membrane protein"	
AUTHORS		/codon_start=3	
TITLE		/transl_table=11	
JOURNAL		/product="p28"	
MEDLINE		/protein_id="AAAG14355.1"	
PUBMED		/db_xref="GI:10181083"	
REFERENCE		/translation="SIISHSVKNLPVQVNTLEEVPRVTSIAITLDIGYLGGEIGIR	
AUTHORS		EIF"	
TITLE		170. .1051	
JOURNAL		/gene="p28-1"	
MEDLINE		170. .1051	
PUBMED		/gene="p28-1"	
REFERENCE		/note="outer membrane protein"	
AUTHORS		/codon_start=1	
TITLE		/transl_table=11	
JOURNAL		/product="p28-1"	
MEDLINE		/protein_id="AAAG14356.1"	
PUBMED		/db_xref="GI:10181084"	
REFERENCE		/translation="MNNKLFITINTVLVCLLSLPNISSSKAINNNKYYGLYISGO	
AUTHORS		YKPSVSFVSNSVKETNVTKNLIALKKDDVDSIEKTDAISVGSINPSNPTIPYTAFFQ	
TITLE		PKRVNSIFHTVMRNDGLSIISVINVCYDFSLNLSISIPYICGAGVDAIEFDFVL	
JOURNAL		HIKAYQSKLGIAYSLPSILFASLYHKVMGNQFNKLVQHVAVELASIPKITSAVA	
MEDLINE		TLNIGYFGGEIGEARLTF"	

Db	269	AAAAAGTACTACGGATTATATATACAGTGGACAATAATAAACCCAGTGTCTCTGTTTTCAGT	328
Qy	49	LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp	68
Db	329	AATTTTTCAGTTAAAGAACCAATGTCATAACTTAAACCTTATAGCTTTAAAAAAGAT	388
Qy	69	Leu-----LeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPhe	83
Db	389	GTTGACTCTATTGAAACCAAGACTGATGCGAGTGTAGTATTAGTAACCCATCAAAATTT	448
Qy	84	AsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeuGlyPheSerGlyIlePheGly	103
Db	449	ACTATCCCTTACAGCTGATTTTCAAGATAATCTGCTCAATTTCAATGAACTATGTTG	508
Qy	104	TyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIle	123
Db	509	TACACCTTGTCTGAAGGTACAAGAGTTGAATAGAGGTTCTTATGAGGAATTTGATGT	568
Qy	124	LysAsp---AsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGlu	142
Db	569	AAAAACCTGGAGCTATACACTAAGTGATGCTCTCGCTATTTTCATTAGCACGTGAA	628
Qy	143	IleSerGlyClySerAsnAsnPro-----AlaAsnAsnLysTyrValThrLeu	158
Db	629	ATGAAAGGTAAATAGTTTACACCTAAGAAAAAGTTTCTAATAGTATTTTTCACACTGA	688
Qy	159	Ile---AsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValAsp	177
Db	689	ATGAGAAATGATGATTAATCTATAATATCTGTTATAGTAATGTTTGCTACGATTCTCT	748
Qy	178	GlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAlaspthrIleAsp	197
Db	749	TTGAACAATTTGTCATATGCGCTTACATATGTTGGAGGACGAGGTAGATGCTATAGAA	808
Qy	198	PheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThr	217
Db	809	TTCTTCGATGATTACACATTAAGTTTGCATATCAAGCAAGCTAGTATTGCTTATTCT	868
Qy	218	ValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLys	237
Db	869	CTACCATCAACATTAAGTCTCTTGTGCTAGTTTATATACATAAGTAATGGCAATCAA	928
Qy	238	PheGlyLysIleProValAsnTyrProCysAspTyrProSerProThrProAsnSer	257
Db	929	TTTAAAAATTTAAATGTCCAACATGTTGCTGAACCTTGCA-----AGT	970
Qy	258	LysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySer	277
Db	971	ATACCTAAAAATT---ACATCCGCGAGTTGCTACACTTAATATTGTTATTGAGGTTGAA	1027
Qy	278	IleGlyIleLys 281	
Db	1028	ATTGGTGCAAGA 1039	
RESULT 5			
AF324792			
LOCUS	AF324792	6913 bp	DNA linear BCT 11-APR-2001
DEFINITION	Ehrlichia canis phosphoribosylaminoimidazole carboxylase (purk) gene, complete cds; major outer membrane protein gene cluster 2, complete sequence; and u6 gene, partial cds.		
ACCESSION	AF324792		
VERSION	AF324792.1	GI:13591681	
KEYWORDS	Ehrlichia canis.		
SOURCE	Ehrlichia canis.		
ORGANISM	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales; Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.		
REFERENCE	1 (bases 1 to 6913)		
AUTHORS	Ohashi,N., Unver,A., Zhi,N. and Rikihisa,Y.		
TITLE	Cloning and characterization of multigenes encoding the immunodominant 30-kilodalton major outer membrane proteins of Ehrlichia canis and application of the recombinant protein for serodiagnosis		

JOURNAL	J. Clin. Microbiol.	36 (9),	2671-2680 (1998)
MEDLINE	983711112		
PUBMED	9705412		
REFERENCE	2 (bases 1 to 6913)		
AUTHORS	Ohashi,N., Rikihisa,Y. and Unver,A.		
TITLE	Analysis of transcriptionally Active Gene Clusters of Major Outer Membrane Protein Multigene Family in Ehrlichia canis and E. chaffeensis		
JOURNAL	Infect. Immun.	69 (4),	2083-2091 (2001)
MEDLINE	21153566		
PUBMED	11254561		
REFERENCE	3 (bases 1 to 6913)		
AUTHORS	Ohashi,N., Rikihisa,Y. and Unver,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-NOV-2000)	Veterinary Biosciences, Ohio State University	1925 Coffey Rd., Columbus, OH 43210, USA
FEATURES	Location/Qualifiers		
source	1..6913		
	/organism="Ehrlichia canis"		
	/strain="Oklahoma"		
	/db_xref="taxon:944"		
gene	329..1420		
	/gene="purk"		
CDS	329..1420		
	/gene="purk"		
	/codon_start=1		
	/transl_table=11		
	/product="phosphoribosylaminoimidazole carboxylase"		
	/protein_id="AAK31311.1"		
	/db_xref="GI:13591682"		
	/translation="MLHNSYVSPGSTIGIGGQLGKMISIAATNLGYKTHLLTDNQD SPINVTNVTILDNYKELLEFPASSVDIATLEPENIPSTPIDLSOKTKVYPGKT ALHISONRIREKQYIRNLGIKTANFRIIONYDNLVHNLVLELGYPALLKRTTELGYDCKG QYLITQPCDLNTLSLWQOEYILEKFVKIHKELISLISKSINGSIEFFPPVAENHFTN GLATSSVPASISEEINTQAKNLQIAESFNVLGILAEFFITDQELVNDIAPRP HNSIYTYGKTKASKNRKMGHINRLNKHEI"		
gene	1756..2598		
	/gene="p30-10"		
CDS	1756..2598		
	/gene="p30-10"		
	/note="P30 family"		
	/codon_start=1		
	/transl_table=11		
	/product="major outer membrane protein p30-10"		
	/protein_id="AAK31312.1"		
	/db_xref="GI:13591683"		
	/translation="MNYKKILYRSALISLMSILPYOSFADPVGSRNTDNKGEFYISAK YNPSISHFRKFAEETPIINGTNSLTKKVFLKKGDDITKKDDFTRVAPGIDFQNNLIS GFSGSIGYSMDGPRIEAAAYQOFNPKNTDNDTNGEYIKHFALSRKMDAQOQYVY LKNDGITFMSLMVNTCYDITAEQSVFVACAGIGADLITIFKDLNLKPAYOCKIGIS YPITPEVSAFIGGYHGVIGNKFEPITVTPVVLNDAPQTTASVTLVDVGYFGGEIGM RETF"		
gene	2926..3756		
	/gene="p30-4"		
CDS	2926..3756		
	/gene="p30-4"		
	/note="P30 family"		
	/codon_start=1		
	/transl_table=11		
	/product="major outer membrane protein p30-4"		
	/protein_id="AAK31313.1"		
	/db_xref="GI:13591684"		
	/translation="MNCCKILITTLVSLTILLPGISFSKPTHENNTCGNFYLGKYY PSISHFGNFSAKEERTTIGIFGLKESWTGGKILDKREHAFNIPNYSFTENNPFLGF AGVIGYSIGSPRIEVEYETDVQNPQDGFKNDAHKYCALNSDSKTKMSKGFVLK NEGLSDISLMNVYDIINKRMPFSYICAGIGTDLIFMFDATINHKAAYQGLGFNYP ISPEANISMGVHFHKVTNNEFRVPVLLTAGGLAPONLFAIVKLSICHGFLGFGYRVSE "		
gene	3967..4830		
	/gene="p30a"		
CDS	3967..4830		
	/gene="p30a"		

[illegible]

Alignment Scores:
Pred. No.: 7.67e-29 Length: 3538
Score: 424.00 Matches: 98
Percent Similarity: 50.87% Conservative: 49
Best Local Similarity: 33.91% Mismatches: 102
Query Match: 28.08% Indels: 40
DB: 1 Gaps: 10

US-09-846-808-1 (1-284) x AF125277 (1-3538)

Qy 13 ValMetLeuLeuPheThrProHisLeuAlaSerValLeuAsnAspHisAsn 32
Db 1027 ATCTCGTTAATGTCATTTTACCATATACAGTCTTTTCGAGAACCTGTAAAGTCAATAAC 1086
Qy 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArg 44
Db 1087 ATTGGAATGAAATGCTAAAGAGGGTCTACATTAAGTGAATAACACCCAGGCATA 1146
Qy 45 GlnHisLeuSerLysLeuLeuLysGluSerAla-----AlaAsnThr 59
Db 1147 CCACACTTCAGAAATTTTCTGCTGAGGAAACTCTCTGTATACGGTAAAGACTCTCAACT 1206
Qy 60 ValGluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAsp 79
Db 1207 AAAAGGTTATTTGGGTTAAAAAGGAG-----GGTTCTATAACAAATAACAGTGTAT 1257
Qy 80 AsnThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeu--GlyPhe 98
Db 1258 TTCACTAGACAGATATATCGTTGAGGCG-----CAAAATATTTTATCTCAGGTTTC 1311
Qy 99 SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 1312 TCAGGAACATAGTTATATCATGAT--GGACCAAGAGTAGATGAAGTCGCATAC 1368
Qy 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
Db 1369 CAAAATTCACACCCAAATAATCCAGCTAATGAACTAGTACTAGTATTACTATAAACAC 1428
Qy 137 PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnLysTyrVal 156
Db 1429 TATGATTATCTCGT-----GCAGAGCTATGGCAGATAAAAATATGTT 1473
Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuAsnValCysTyrAspVal 176
Db 1474 GTACTTACAAATAATGGAGTAACTTTTCATCAATTAATGTTCAATGCTGCTATGACATT 1533
Qy 177 AspGlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIle 196
Db 1534 ACAGCAGAAGGAGTACCTTTCATTCATATCATGCTGCTGCTGTTGGTGCAGATCTTATA 1593
Qy 197 AspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyr 216
Db 1594 TCTATATTGATGATATAAATTTAAATTTTCTTACCAGGTAATAATGGTATTAGTTAT 1653
Qy 217 ThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLys 236
Db 1654 CCAATTACTCTGAAATTTCTGCATTTATTCGCGGATATTACCATGGAGTAATAGGTAAC 1713
Qy 237 LysPheGluLysIleProValAsnTyrPro-----CysAspTyrProSerProThr 253
Db 1714 AAGTATAACAAGTACCTGTAAAGCTTCTGTAACCTTTTAACAGATGCTCTCTCAA----- 1767
Qy 254 ProProAsnSerLysProHisValHisThrAlaLeuAlaMetLeuSerIleGlyTyr 273
Db 1768 -----AGCACTTCCGCTTCAGTAACCTCTTTGACCGCTGGATAT 1803
Qy 274 TyrGlyGlySerIleGlyIleLysPhe 282
Db 1804 TTTGGTGGTGAACCTTGGAGTAAGGTTT 1830

RESULT 8
AF125279

LOCUS AF125279 3541 bp DNA linear BCT 20-APR-1999
DEFINITION Cowdria ruminantium isolate Antigua major antigenic protein 1 like
protein and major antigenic protein 1 (MAP1) genes, complete cds.
ACCESSION AF125279
VERSION AF125279.1 GI:4589109
KEYWORDS heartwater rickettsia.
SOURCE Cowdria ruminantium
ORGANISM Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Cowdria.
REFERENCE 1 (bases 1 to 3541)
AUTHORS Sulsona,C.R., Mahan,S.M. and Barbet,A.F.
TITLE The map1 Gene of Cowdria ruminantium Is a Member of a Multigene
Family Containing Both Conserved and Variable Genes
JOURNAL Biochem. Biophys. Res. Commun. 257 (2), 300-305 (1999)
REFERENCE 2 (bases 1 to 3541)
AUTHORS Sulsona,C.R., Mahan,S.M. and Barbet,A.F.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-1999) College of Veterinary Medicine Department
of Pathobiology, University of Florida, P.O. Box 110880,
Gainesville, FL 32610, USA
FEATURES
source Location/Qualifiers
1..3541
/organism="Cowdria ruminantium"
/isolate="Antigua"
/db_xref="taxon:779"
989..1837
/note="ORE2"
/codon_start=1
/transl_table=11
/product="major antigenic protein 1 like protein"
/protein_id="AAD26353.1"
/db_xref="GI:4589110"
/translation="MNKKILVRSALISLMSFLPYQSPAEVPSVSSNNIGNNAKEGFYI
SAKNPSIHPKFSAEPTVYKDSPTKKVFKLKEGSIKISDFTIDISFGQNN
FISFGSIGYIMDPRVEIEAAYOKPNPNPANEPTSDYIKHYGSRAMADKKY
VVLNNGVTSSLMFNACYDITAEGVFPIYACAGIGADLISIFDDINLKFAYQKIG
ISVPTTPEISAFIGGYGVHGVGNKYNKVPKLVLTLDPOSTASVTLDAFYGGEL
GVRFYF"
2210..3055
/gene="MAP1"
2210..3055
/gene="MAP1"
/codon_start=1
/transl_table=11
/product="major antigenic protein 1"
/protein_id="AAD26354.1"
/db_xref="GI:4589111"
/translation="MNCKKIFITSLISLVSFLPQVSDVIOEDSPVGSVYISAKY
MPTASHFGKMSIKEDSRDTKVFGKKDWGKVTSSNTIFTEKDYSPKYNNPFLGF
AGATGYSMNGPRIEFESYETFDVKNPGNKNDAHMYCALDTSSTGGAASVVMV
KNEMLTISLMLNACYDMLDMPVSPVCAGICTDLVSLNSTPNKLSYOGKLGISY
SINPEASIFIGGHRFVIGNEFKDITTSKIFNTSGTATPGFASAILDVCHFGIEIG
GRFYF"
BASE COUNT 1282 a 451 c 479 g 1329 t
ORIGIN
Alignment Scores:
Pred. No.: 7.68e-29 Length: 3541
Score: 424.00 Matches: 98
Percent Similarity: 50.87% Conservative: 49
Best Local Similarity: 33.91% Mismatches: 102
Query Match: 28.08% Indels: 40
DB: 1 Gaps: 10
US-09-846-808-1 (1-284) x AF125279 (1-3541)
Qy 13 ValMetLeuLeuPheThrProHisLeuAlaSerValLeuAsnAspHisAsn 32
Db 1025 ATCTCAATGTCATTTTACCATATACAGTCTTTTCGAGAACCTGTAAAGTCAATAAC 1084
Qy 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArg 44
Db 1025 ATCTCAATGTCATTTTACCATATACAGTCTTTTCGAGAACCTGTAAAGTCAATAAC 1084


```
Db 1085 ATTGGAATGAAATGCTAAAGAGGTTCTACATAAGTGCAAAATACAAACCAAGCATA 1144
Qy 45 GlnHisLeuSerLysLeuLeuLeuLysGluSerAla-----AlaAsnThr 59
Db 1145 CCACACTTCAGAAAATTTCTGCTGAGGAACCTCTGTATACGGTAAAGACTCTCCAAC 1204
Qy 60 ValGluValPheGlyLeuLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAsp 79
Db 1205 AAAAAGGTATTGGGTTAAAAAGGAG-----GGTCTATATACAAAATACAGTGAT 1255
Qy 80 AsnThrAsnPheAsnLeuLysTyrAsnProTyrTyrGluAsnAsnArgLeu---GlyPhe 98
Db 1256 TTCCTAGAACAGATATATCGTTGAGGCG-----CAAAATATTTATCTCAGGTTTC 1309
Qy 99 SerClyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 1310 TCAGGAAGCATAGGTATATCATGGAT---GGACCAAGAGTAGAGATTGAAGCTGCATAC 1366
Qy 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
Db 1367 CAAAAATCAACCAAAATCCAGCTAATGAACCTGATAGTACTAGTACTATAAACAC 1426
Qy 137 PheAlaLeuAlaLysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
Db 1427 TATGGATTATCTCGT-----GCAGAAGCTATGGCAGATAAAAAATATGTT 1471
Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
Db 1472 GTACTTACAATAANTGGAGTAACCTTTTCATCATTAATTAATGTTCAATGCTGCTATGACATT 1531
Qy 177 AspGlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIle 196
Db 1532 ACAGCAGAAGGACTACCTTTCATCCATATGCATGCTGGTATTGGTCAGATCTATATA 1591
Qy 197 AspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyr 216
Db 1592 TCTATATTTGATGATATAAATTTAAATTCGCTTACCAAGGTAATAATTTGTTATAGTTAT 1651
Qy 217 ThrValSerProGlnValSerValPheIleGluGlyTyrThrHisGlyLeuPheGlyLys 236
Db 1652 CCAATTACTCCTGAATTTCTGCATTATTGGCGGATATTACCATGAGTAATAGGTAAC 1711
Qy 237 LysPheGluLysIleProValAsnTyrPro-----CysAspTyrProSerProThr 253
Db 1712 AAGTATAACAAAGTACCTGTAAGACTTCCTGTAACCTTTACACAGATGCTCCTCAA----- 1765
Qy 254 ProProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyr 273
Db 1766 -----AGCACTTCGCTTACGTAACCTCTTGACGCTGGATAT 1801
Qy 274 TyrGlyGlySerIleGlyIleLysPhe 282
Db 1802 TTGGGCGTGAACCTTGAGTAAGGTTT 1828

RESULT 9
AF125278 3572 bp DNA linear BCT 20-APR-1999
LOCUS Cowdria ruminantium isolate Umbanein major antigenic protein 1 like
DEFINITION protein and major antigenic protein 1 (MAP1) genes, complete cds.
ACCESSION AF125278
VERSION AF125278.1 GI:4589106
KEYWORDS heartwater rickettsia.
SOURCE Cowdria ruminantium
ORGANISM Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Cowdria.
REFERENCE 1 (bases 1 to 3572)
AUTHORS Sulsona,C.R., Mahan,S.M. and Barbet,A.F.
TITLE The map1 Gene of Cowdria ruminantium Is a Member of a Multigene
Family Containing Both Conserved and Variable Genes
JOURNAL Biochem. Biophys. Res. Commun. 257 (2), 300-305 (1999)
REFERENCE 2 (bases 1 to 3572)
AUTHORS Sulsona,C.R., Mahan,S.M. and Barbet,A.F.
```

```
TITLE Direct Submission
JOURNAL Submitted (02-FEB-1999) College of Veterinary Medicine Department
of Pathobiology, University of Florida, P.O. Box 110880,
Gainesville, FL 32610, USA
FEATURES
source 1..3572
/organism="Cowdria ruminantium"
/isolate="Umbanein"
/db_xref="taxon:779"
989..1837
/notes="ORF2"
/codon_start=1
/transl_table=11
/product="major antigenic protein 1 like protein"
/protein_id="AAD26351.1"
/db_xref="GI:4589107"
SKYNPISPIHFRKFSAEETPVYKSDTKKKVFKLKEGSIYKYSIDTRTDISEGQNN
FISGSGSIGYIMDPRVEIARAYQFNKPNANETDTSYKHYGLISRAEMADKKY
VLTNNGVTFSSLMFNACVDITAEQVFPFIPYACAGIGADLISIFDDINLKFAVQGGKIG
ISYPTPEISAFIGYHYGVIGNKYNKVPKLPVTLTDPAQSTISASVTLDAFYEGGEL
GVRFTE"
gene 2210..3082
/genes="MAP1"
CDS 2210..3082
/genes="MAP1"
/codon_start=1
/transl_table=11
/product="major antigenic protein 1"
/protein_id="AAD26352.1"
/db_xref="GI:4589108"
/transl_table="MCKKIFITSLISLVFLPVSFSDVIOEDSNPAGSVVISARY
PFTASHFGKMSIKEDSKNTQTVFGLKDDGVKTPSSDSCSNSIITEKDYSPKYENN
MELGPAAGATGYSMNGPRIEFESYEDFYKNGCNKNDAMHWCALDTPGPGTQGNAT
LNSVVMKNELTDIALMLNACVDITAEQVFPFIPYACAGIGADLISIFDDINLKFAVQGGKIG
KGLGISYINPEASIFIGGHRVIGNEFKDITTSKIFTSTGLKATAAASPGFASATLD
VCFHGEIIGGVFV"
BASE COUNT 1303 a 472 c 476 g 1321 t
ORIGIN
Alignment Scores: 7.76e-29 Length: 3572
Pred. No.: 424.00 Matches: 98
Score: 50.87% Conservative: 49
Percent Similarity: 33.91% Mismatches: 102
Best Local Similarity: 28.08% Indels: 40
Query Match: 1 Gaps: 10
DB:
US-09-846-808-1 (1-284) x AF125278 (1-3572)
Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
Db 1025 ATCTCATTAATGTCATTTTACCATATCAGTCTTTTGCAGAACCTGTAAGTTCAANTAAC 1084
Qy 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArg 44
Db 1085 ATTGGAATGAAATGCTAAAGAGGTTCTACATAAGTGCAAAATACAAACCAAGCATA 1144
Qy 45 GlnHisLeuSerLysLeuLeuLysGluSerAla-----AlaAsnThr 59
Db 1145 CCACACTTCAGAAAATTTTCTGCTGAGGAACCTCTGTATACGGTAAAGACTCTCCAAC 1204
Qy 60 ValGluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAsp 79
Db 1205 AAAAAGGTATTGGGTTAAAAAGGAG-----GGTCTATATACAAAATACAGTGAT 1255
Qy 80 AsnThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeu---GlyPhe 98
Db 1256 TTCCTAGAACAGACATATCGTTGAGGCG-----CAAAATATTTATCTCAGGTTTC 1309
Qy 99 SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 1310 TCAGGAAGCATAGGTTTATATCATGGAT---GGACCAAGAGTAGAGATTGAAGCTGCATAC 1366
```


QY 197 AspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyr 216
 Db 1551 TCTATATTTGATGATATAATAATTAAATTTGCTTACCAGGTAATAATTGGTATTAGTTAT 1610
 QY 217 ThrValSerProGlnValSerValPheIleGluGlyTyrThrHisGlyLeuPheGlyLys 236
 Db 1611 CCAATTACTCCTGAAATTTCTGCATTTATTTGGCGGATATTACCATGGAGTAATAGGTAAC 1670
 QY 237 LysPheGluLysIleProValAsnTyrPro-----CysAspTyrProSerProThr 253
 Db 1671 AAGTATAACAAGTACCTGTAAGCTTCTTAACGTTTAAACAGATCTCTCTCAA----- 1724
 QY 254 ProProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyr 273
 Db 1725 -----AGCACTTCGGCTTCAGTAACTCTTCGACGCTGGATAT 1760
 QY 274 TyrGlyGlySerIleGlyLysPhe 282
 Db 1761 TTTGGTGGTGAACCTTGAGTAAGGTTT 1787
 RESULT 11
 LOCUS AF125275 3551 bp DNA linear BCT 20-APR-1999
 DEFINITION Cowdria ruminantium isolate Crystal Springs major antigenic protein
 1 like protein and major antigenic protein 1 (MAP1) genes, complete
 cds.
 ACCESSION AF125275
 VERSION AF125275.1 GI:4589097
 KEYWORDS heartwater rickettsia.
 SOURCE Cowdria ruminantium.
 ORGANISM Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 Rickettsiaceae; Ehrlichieae; Cowdria.
 1 (bases 1 to 3551)
 Sulsona,C.R., Mahan,S.M. and Barbet,A.F.
 The map1 Gene of Cowdria ruminantium Is a Member of a Multigene
 Family Containing Both Conserved and Variable Genes
 Biochem. Biophys. Res. Commun. 257 (2), 300-305 (1999)
 2 (bases 1 to 3551)
 Sulsona,C.R., Mahan,S.M. and Barbet,A.F.
 Direct Submission
 Submitted (02-FEB-1999) College of Veterinary Medicine Department
 of Pathobiology, University of Florida, P.O. Box 110880,
 Gainesville, FL 32610, USA
 FEATURES
 Location/Qualifiers
 1..3551
 /organism="Cowdria ruminantium"
 /isolate="Crystal Springs"
 /db_xref="taxon:779"
 991..1839
 /note="ORF2"
 /codon_start=1
 /transl_table=11
 /product="major antigenic protein 1 like protein"
 /protein_id="AAD26345.1"
 /db_xref="GI:4589098"
 /translation="MNKKILVRSALISLMSFLPYQSFAPVSSNNIGNENAKEFYI
 SAKNPSIPHRKFAEETPVIGKDSKPKVFLAKESGIKYSIDFTRIDISFEQNN
 FISGSGIGIMDGPVEIEAAYQKPNKPANETIDTSYKKHYGLSRAETMTDKY
 VLTNNGVTSLSMFNACDITAEGVPIPYACAGIGADLISIFDDIILKRAYQKIG
 ISYPTITPISAFIGYHGVIGNKYNKVPKLPVLTDAPOSSTSASVTLDAFYFGEL
 GVRFTF"
 2199..3062
 /gene="MAP1"
 2199..3062
 /gene="MAP1"
 /codon_start=1
 /transl_table=11
 /product="major antigenic protein 1"
 /protein_id="AAD26346.1"
 /db_xref="GI:4589099"
 /translation="MNCKKIFITSLISVSLPGVSPFSDVIQEDSNPAGSVYISAKY
 MPTASHFKMSIKEDSKNTQTVEGLKQDWGVKTPSDSSNSTNITFTKDYSFYENN

PFLGFAIGYSMNGPRIEFVSYETFDVKPNPGNYKNNAHMYCALDTAAQNSTNGAG
 LTKTSVMKNEINTNISLMLNACYDIPVSPVYACIGIGTDLVSVINATPNKLSYO
 CKLGISYSINSEASIFIGGHFRHVRIGNEEKDIATLKIFTSKGTGINSNPGFASATLDVCH
 FGIEIGGRFVF"
 BASE COUNT 1308 a 477 c 464 g 1302 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,17e-28 Length: 3551
 Score: 422.00 Matches: 97
 Percent Similarity: 50.52% Conservative: 49
 Best Local Similarity: 33.56% Mismatches: 103
 Query Match: 27.95% Indels: 27 40
 DB: 1 Gaps: 10
 US-09-846-808-1 (1-284) x AF125275 (1-3551)
 QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
 Db 1027 ATCTCATTAATGTCAATTTTACCATATCAGTCTTTTGCAGAACCTGTAGTTCACAAATAAC 1086
 QY 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArg 44
 Db 1087 ATTGSAATCAAATGCTAAAGAGGGTTCTACATAAGTGCACAAATACAAACCAAGCAT 1146
 QY 45 GlnHisLeuSerLysLeuLeuLysGluSerAla-----AlaAsnThr 59
 Db 1147 CCNACTTTCAGAAAATTTTCTGCTGAGGAACCTCGTATACGGTAAGACTCTCCAACT 1206
 QY 60 ValGluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAsp 79
 Db 1207 AAAAGGTATTTGGTTAAAAAGGAG-----GGTCTATAACAAAAATACAGTGAT 1257
 QY 80 AsnThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeu---GlyPhe 98
 Db 1258 TTCACGTAGAACATATATCGTTTGAGGCG-----CAAATAATTTATCTCAGGTTTC 1311
 QY 99 SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
 Db 1312 TCAGGAGCATAGTTATATCATGGAT---GGACCAAGAGTAGAGATTGAAGCTGCATAC 1368
 QY 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
 Db 1369 CAAAAATTCAACCCCAAAAAATCCAGCTAATGAAACTGATACTAGTGATTACTATAAACAC 1428
 QY 137 PheAlaLeuAlaLysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
 Db 1429 TATGGATTATCTCGT-----GCAGAAACCATGACAGATAAAAAATATGTT 1473
 QY 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
 Db 1474 GTACTTACAATANTGAGTAACATTTTTCATCATTAATGTTTAAATGCTGCTATGACATT 1533
 QY 177 AspGlyLeuLysHisAsnIleThrThrTyrSerCysLeuGlyPheGlyValAspThrIle 196
 Db 1534 ACAGCAGAGGAGTACCTTTTCATTCATATGTCATGCTGCTGTTTGGTCAGATCTTATA 1593
 QY 197 AspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyr 216
 Db 1594 TCTATATTTGATGATATAAATTTTAAATTTTGGCTTACCAGGTAAATTTGTTATAGTTAT 1653
 QY 217 ThrValSerProGlnValSerValPheIleGluGlyTyrThrHisGlyLeuPheGlyLys 236
 Db 1654 CCAATTACTCTGAAATTTTTCGATTTTATGGCGGATATTACCATTGAGTAAGTAGTAAC 1713
 QY 237 LysPheGluLysIleProValAsnTyrPro-----CysAspTyrProSerProThr 253
 Db 1714 AAGTATAACAAGTACCTGTAAAGCTTCTGTAACTTTTAAACAGATCTCTCTCAA----- 1767
 QY 254 ProProAsnSerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyr 273
 Db 1768 -----AGCACTTCGGCTTCAGTAACTCTTCGACGCTGGATAT 1803

```

QY 274 TyrGlyGlySerIleGlyIleLysPhe 282
Db 1804 TTTGGTGGTGAACCTGGAGTAAGGTTT 1830

RESULT 12
AX042314 AX042314 843 bp DNA linear PAT 23-NOV-2000
LOCUS
DEFINITION Sequence 10 from Patent WO0065063.
ACCESSION AX042314
VERSION AX042314.1 GI:11340974
KEYWORDS
SOURCE
ORGANISM Ehrlichia chaffeensis.
Bacteria; Ehrlichia chaffeensis
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichiae; Ehrlichia; canis group.
REFERENCE
1 (bases 1 to 843)
AUTHORS Barbet,A.F., Bowie,M.V., Ganta,R.R., Burrigde,M.J., Mahan,S.M.,
Mcquire,T.C., Rurangirwa,F.R., Moreland,A.L., Simbl,B.H.,
Whitmire,W.W. and Alleman,A.R.
TITLE Nucleic acid vaccines against rickettsial diseases and methods of
use
JOURNAL Patent: WO 0065063-A 10 02-NOV-2000;
UNIVERSITY OF FLORIDA (US)
FEATURES
source Location/Qualifiers
1. .843
/organism="Ehrlichia chaffeensis"
/db_xref="taxon:945"
BASE COUNT 282 a 127 c 159 g 275 t
ORIGIN

Alignment Scores:
Pred. No.: 1.81e-28 Length: 843
Score: 411.50 Matches: 103
Percent Similarity: 52.08% Conservative: 47
Best Local Similarity: 35.76% Mismatches: 113
Query Match: 27.25% Indels: 25
DB: 6 Gaps: 11

US-09-846-808-1 (1-284) x AX042314 (1-843)
QY 7 ArgLysPheValLeu-----TrpValMetLeuIleLeuPheThrProHisIleSer 23
Db 12 AAAAAATTTTATAACAACTACATTAGTAGCTGCTAATGCTCTCTTACCTGGAAATACA 71
QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
Db 72 TTTTCTGATGAGTACAGAACCAATGTTGCTGGTAATTTCTATATATCATGCGGAAATAT 131
QY 41 LysProAlaArgGlnHisLeuSerLysLeuIleLysGluSerAlaAlaAsnThrVal 60
Db 132 GTACCAAGTGTTCACATTTTGGCGTATTCCTGCTAAACAGGAAGAAATACACAAATC 191
QY 61 GluValPheGlyLeuLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
Db 192 GGAGTATTTGGATTAAACAGATGGGATGGCAGCACAAATATCTAAAAATCTCCAGAA 251
QY 81 ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
Db 252 AATACATTTAAGTTCACAAATTTATTCATTAAATATGAAATAATATCCATTTCTAGGTTT 311
QY 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 312 GCAGGAGCTGTGGTATTTAATGAAT---GGTCCAGAAATAGAGATAGAAATGTCCAT 368
QY 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
Db 369 GAACATTTGATGGAACAAACAGGTAATAACTATAAGAAC---GATGCTCACAAATAT 425
QY 137 PheAlaLeuAlaLysGluIleSerGlySerAsnProAlaAsnAspLysTyrVal 156
Db 426 TATGCTTTAACCCATMAC--AGTGGGGGAAGCTAAGCAATGACAGGTGATAGTTGTT 482
QY 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176

```

```

Db 483 TTTCTAAAAAATGAAGGACTACTTGATATATCACTTATGTTGAATGCATGCTATGATGA 542
QY 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValasp 194
Db 543 ATAAGTGAAGGAATACCTTCTCT-----CCTTACATATGTCAGGCTTGTGACTGAT 596
QY 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
Db 597 TTAATATCCATGTTTGAAGCTATAAACCCATAAAATTTCTTATCAAGGAAGTTAGGTTTG 656
QY 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
Db 657 AGTTACTCCATAAGCCAGCAAGCTCTGTTTGTGGTGGACATTTTCATAGGTGATA 716
QY 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro 254
Db 717 GGAATGAATTCAGA-----GATATCTCTCTATGATACCC 752
QY 255 ProAsnSerLysProHisValHisThrAlaLeuAlaMetLeuSerIleGlyTyrTyr 274
Db 753 AGTACCTCAACTCTCACAGGTAATCACTTACTATAGTAACACTAAGTGTATGCCACTTT 812
QY 275 GlyGlySerIleGlyIleLysPhe 282
Db 813 GGAGTGGAACTTGGAGGAAGGTTT 836

RESULT 13
AF062761
LOCUS AF062761 4683 bp DNA linear BCT 18-JUL-1998
DEFINITION Ehrlichia chaffeensis 28 kDa major surface antigen multi-gene
locus, partial sequence.
ACCESSION AF062761
VERSION AF062761.1 GI:3327958
KEYWORDS Ehrlichia chaffeensis.
SOURCE Ehrlichia chaffeensis.
ORGANISM Ehrlichia chaffeensis
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichiae; Ehrlichia; canis group.
REFERENCE
1 (bases 1 to 4683)
AUTHORS Reddy,G.R., Sulsona,C.R., Barbet,A.F., Mahan,S.M., Burrigde,M.J.
and Alleman,A.R.
TITLE Molecular characterization of a 28 kDa surface antigen gene family
of the tribe Ehrlichiae
JOURNAL Biochem. Biophys. Res. Commun. 247 (3), 636-643 (1998)
MEDLINE 98321180
REFERENCE
2 (bases 1 to 4683)
AUTHORS Reddy,G.R., Sulsona,C.R., Barbet,A.F., Mahan,S.M., Burrigde,M.J.
and Alleman,A.R.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-1998) Diagnostic Medicine/Pathobiology, College
of Veterinary Medicine, Kansas State University, 1800 Denison Ave,
Manhattan, KS 66506, USA
FEATURES
source Location/Qualifiers
1. .4683
/organism="Ehrlichia chaffeensis"
/strain="Arkansas"
/db_xref="taxon:945"
<1--134
/feature="similar to Cowdria ruminantium MAP1 gene; ORF 1;
silent copy"
/codon_start=3
/transl_table=11
/evidence=not_experimental
/product="28 kDa major surface antigen-1"
/protein_id="AAC26717.1"
/db_xref="GI:3327960"
/translation="NEFRDISTLKAFATPPSSAATPDLATVTLVSCHFGVGLGRFNF"
179. .188
/feature="palindromic sequence part 1"
/evidence=not_experimental
203. .212
/feature="palindromic sequence part 2"

```

```
misc_feature
-35_signal
-10_signal
RBS
CDS

/evidence=not_experimental
226..239
/note="g-rich sequence"
349..354
/evidence=not_experimental
375..380
/evidence=not_experimental
430..434
/evidence=not_experimental
439..1299
/note="similar to Cowdria ruminantium MAP1 gene; ORF 2;
silent copy"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="28 kDa major surface antigen-2"
/db_xref="GI:3327961"
/translation="MNCCKFFITTTALVSLMFLPGISLSDPVQDDNISGNFYISKYK
PSASHFGVFSKAEKERTTVGVFGIQDWDRCVISRITLSDIFTVPNYSFKYENNFPG
FAGAIQYSDGPRIELEVSYEAFDVKNOGNNYKNEAHRYALSHLLGTETQIDGAGSA
SVFLNELDKSFMNACVDVISEGIPFSPYICAGIGIDLVSMEFANPKISYOGKLSY
GLSYPISPEASVFIGGHFKVIGNEFRDIPITPMSALAGNIPAIIVLDDVFYFGI
ELGGRNFQL"
1300..1309
/note="palindromic sequence part 1"
/evidence=not_experimental
1324..1333
/note="palindromic sequence part 2"
/evidence=not_experimental
1347..1361
/note="g-rich sequence"
1473..1479
/evidence=not_experimental
1498..1504
/evidence=not_experimental
1554..1558
/evidence=not_experimental
1563..2399
/note="similar to Cowdria ruminantium MAP1 gene; ORF 3;
silent copy"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="28 kDa major surface antigen-3"
/db_xref="GI:3327962"
/translation="MNCCKFFITTTALVSLMFLPGISLSDPVQDDNISGNFYISKYK
PSASHFGVFSKAEKERTTVGVFGIQDWDREGISSSHNDNHFNNKYSFKYENNFPGF
AGAIGYSDGPRIELEVSYEAFDVKNOGNNYKNEAHRYALSHLLGTETQIDGAGSA
KSEGLDLSFMNACVDIINESIPLSPIYICAGVGTDLISMEFATNPKISYOGKLGLSY
SINPASVFIGGHFKVIGNEFRDIPITLKAFYTSSATPDLAIVTLVSLVCHFGIELGGRF
NP"
2444..2456
/note="palindromic sequence part 1"
/evidence=not_experimental
2465..2477
/note="palindromic sequence part 2"
/evidence=not_experimental
2491..2505
/note="g-rich sequence"
2618..2624
/evidence=not_experimental
2644..2649
/evidence=not_experimental
2699..2703
/evidence=not_experimental
2708..3550
/note="similar to Cowdria ruminantium MAP1 gene; ORF 4;
silent copy"
/codon_start=1
/transl_table=11
/evidence=not_experimental

terminator
terminator
misc_feature
-35_signal
-10_signal
RBS
CDS

/product="28 kDa major surface antigen-4"
/protein_id="AAC26720.1"
/db_xref="GI:3327963"
/translation="MNCCKFFITTTLVSLMFLPGISLSDPVQDDNISGNFYISKYK
PSVSHFGVFSKAEKERTTVGVFGIQDWDGSTISKNSPENTFNPNYSFKYENNFPGF
FAGAVGYLMNGPRIELEVSYEAFDVKNOGNNYKNEAHRYALSHLLGTETQIDGAGSA
FLKNEGLDLSIMNACVDVISEGIPFSPYICAGVGTDLISMEFANPKISYOGKLGSL
SYSISPEASVFIGGHFKVIGNEFRDIPITPMSALAGNIPAIIVLDDVFYFGI
RNF"
3586..3596
/note="palindromic sequence part 1"
/evidence=not_experimental
3610..3622
/note="palindromic sequence part 2"
/evidence=not_experimental
3643..3655
/note="g-rich sequence"
3767..3772
/evidence=not_experimental
3788..3793
/evidence=not_experimental
3847..3851
/evidence=not_experimental
3856..4683
/note="similar to Cowdria ruminantium MAP1 gene; ORF 5;
expressed copy"
/codon_start=1
/transl_table=11
/product="28 kDa major surface antigen-5"
/protein_id="AAC26716.1"
/db_xref="GI:3327959"
/translation="MNYKKVFTTSLISLPGVFSFDPAGSGINGNFYISGKYMP
SASHFGVFSKAEKERTTVGVFGIKONWDGSAISNSPNDVFTVSNYSFKYENNFPGF
AGAIGYSDGPRIELEVSYEAFDVKNOGNNYKNEAHRYALSHLLGTETQIDGAGSA
LKNEGLDLSFMNACVDVIGGEGIPFSPYICAGIGIDLVSMEFATNPKISYOGKLGSL
YSISPEASVFIGGHFKVIGNEFRDIPITPITGSTLAGNYPPIVILDDVFYFGIEMG"
1578 a 739 c 833 g 1533 t
BASE COUNT
ORIGIN

Alignment Scores:
Pred. No.: 1.46e-27 Length: 4683
Score: 411.50 Matches: 103
Percent Similarity: 52.08% Conservative: 47
Best Local Similarity: 35.76% Mismatches: 113
Query Match: 27.25% Indels: 25
DB: 1 Gaps: 11

US-09-846-808-1 (1-284) x AF062761 (1-4683)
QY 7 AtgLYsPheValLeu-----TtpValMetLeuIleLeuPheThrProHisIleSer 23
Db 2717 AAAAAATTTTTTATACAACTACATAGTATCGCTATGCTTCTTCTTACTCGGAATATCA 2776
QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
Db 2777 TTTTCTGATGCAGTACAGAACGACAATGTTGGTGGTAATTTCTATATACAGTGGCAATAT 2836
QY 41 LysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSerAlaAlaAsnThrVal 60
Db 2837 GTACCAAGTGTTTTCACATTTTGGCGTATTTCTGCTAAACAGGAGAAATACACAACAT 2896
QY 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
Db 2897 GGAGTATTTGGATTAAAGCAAGATTTGGGATGGCAGCACATATCTAAAAATTTCTCCAGAA 2956
QY 81 ThrAsnPheAsnIle---LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
Db 2957 AATACATTTAACGTTCCAAATTTATTCATTTAAATATGAAATAATCCATCTTAGGTTTT 3016
QY 99 SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 3017 GCAGGAGCTGTTGGTTATTATTAATGAAT---GGTCCAAGATAGAGTTAGAAATGTCTTAT 3073
```


Qy	260	HisValHisThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGly	279
Db	1165	GCCACTCCAGGCTTTGCATGCACCAATACTTGATGTCGCCATTCGGTATAGAAATGGA	1224
Qy	280	IleLysPheIle	283
Db	1225	GGAAAGGTTGTGA	1236
RESULT 15			
CRU50834			
DEFINITION	CRU50834	1101 bp DNA linear BCT 14-JUL-1996	
DESCRIPTION	Cowdria ruminantium Nyatsanga major antigenic protein 1 (map1)		
ACCESSION	gene, complete cds.		
VERSION	U50834		
KEYWORDS	U50834.1 GI:1418257		
SOURCE	heartwater rickettsia strain-Nyatsanga.		
ORGANISM	Cowdria ruminantium		
REFERENCE	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales; Rickettsiaceae; Ehrlichiaeae; Cowdria.		
AUTHORS	1 (bases 1 to 1101) Reddy,G.R., Sulsona,C.R., Harrison,R.H., Mahan,S.M., Burrridge,M.J. and Barbet,A.F.		
TITLE	Sequence heterogeneity of the major antigenic protein 1 genes from Cowdria ruminantium isolates from different geographical areas		
JOURNAL	Clin. Diagn. Lab. Immunol. 3 (4), 417-422 (1996)		
MEDLINE	96400830		
REFERENCE	2 (bases 1 to 1101) Reddy,G.R.		
AUTHORS	Direct Submission		
TITLE	Submitted (06-MAR-1996) Roman G. Reddy, Pathobiology, University of Florida, 471 Mowry Rd., Gainesville, FL 32611-0880, USA		
JOURNAL	Location/Qualifiers		
FEATURES	1..1101		
source	/organism="Cowdria ruminantium" /strain="Nyatsanga" /db_xref="taxon:779" 208..1080 /genes="map1" 208..1080 /genes="map1" /note="surface protein" /codon_start=1 /transl_table=11 /product="major antigenic protein 1" /protein_id="AAC4145.1" /db_xref="GI:1418258" /translation="MCKKIFITSLISLFLPGVFSVDVQEDSSPAGSVVISARY MPTAFHGKMSIKEDSKNTQVFLKKDWDGVKPTSENSNSTIFTEKDYSPRYEN PFLGAFGRIGYVSNMGRIFEFYSYETFDVKNPGNYKDAHMYCALDTAQPTSNOCAT LASSVMVKNENLTDISLMLNACVDIMLGMPPVSPYVACIGTDLVSVINATPKLQYQ GKLGISYLNPEASIFIGGHFHVINEFKDIATSKIFTSSKASSIPNPGFASATLD VCHFGEIGRFFV"		
BASE COUNT	379 a 158 c 170 g 394 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.18e-26	Length:	1101
Score:	393.00	Matches:	104
Percent Similarity:	53.22%	Conservative:	53
Best Local Similarity:	35.25%	Mismatches:	96
Query Match:	26.03%	Indels:	42
DB:	1	Gaps:	15
US-09-846-808-1 (1-284) x CRU50834 (1-1101)			
Qy	13	ValMetLeuIleLeuPThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn	32
Db	244	ATATCATTAGTGTCATTTACCTGGTGTCCTTTTCTGATGTAATACAGGACAGC	303
Qy	33	-----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer	48

Db	304	AGCCAGCAGCAGCAGTGTTCATCATTTAGCCCAAAATACATGCCAATGCATCACATTTTGGT	363
Qy	49	LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp	68
Db	364	AAATGTCAATCAAGAAGATTCAAAATAATCTCAACACAGTATTTGGTCTAAAAAAGAT	423
Qy	69	LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsn-----	82
Db	424	-----TGGGATGGCGTTAAAGTACCAACATCAGAAAAACAGTAACAATTC	468
Qy	83	-----PheAsnIleLys---TyrAsnProFtyrTyrGluAsnAsnAsnArg---LeuGlyPhe	98
Db	469	ACAAATTTTACTGAAAAGATATCTCTTCAGATATGAAAAAATCCCGTTTTAGGTTTT	528
Qy	99	SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr	118
Db	529	GCTGGACGAATTGGATCTCAATCAAT---GGGCCAAGAATAGAGTTTGAAGTATCTCT	585
Qy	119	GluThrPheHisIleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHis	136
Db	586	GAAACTTTTGTATGCAAAAATCCAGGTGGCAACTATATAAAT---GATGCACACATGTAC	642
Qy	137	PheAlaLeuAlaLysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrVal	156
Db	643	TGTCCTTTA-----GATACACACAGCCCACTAGTAATCAAGGTGCA	684
Qy	157	ThrLeuIle-----AsnAsnGlyIleSerLeuThrSerAlaLeuIle	170
Db	685	ACATTAGCTTCATCTGTATGGTAAAAAATGCAAAATTTACAGATATATCATTAATGTTA	744
Qy	171	AsnValCysTyrAspVal-----AspGlyLeuLysHisAsnIleIleThrTyrSerCys	188
Db	745	AATGCATGTTATGATATAAATGCTTGATGGAATGCCA-----GTTCTCCATATGTATGT	798
Qy	189	LeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyr	208
Db	799	GCAGGTATGGTACTGATTTAGTCTCAGTAATTAATGCTACAAATCCCTAAATTCATAT	858
Qy	209	GlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIleGluGly	228
Db	859	CAAGAAAGCTAGCATAAGTTACTCAATCAATCCTCAAGCTCTATCTTTATTGGTGGGA	918
Qy	229	TyrThrHisGlyLeuPheGlyLysLysPheGlyLysIleProValAsnTyrProCysAsp	248
Db	919	CATTTCATAGATTATAGGTAAATGAATTTAAAGATATTCGTACCTCAAAA-----ATA	972
Qy	249	TyrProSerProThrProProAsnSerLysProHisValHisThrAlaLeuAlaMet	268
Db	973	TTTACTTCAACTAGTAAGCATCATATACCTTAAT---CCTGGCTTTGGCATCAGCAACA	1029
Qy	269	LeuSerIleGlyTyrTyrGlyGlySerIleGlyIleLysPheIle	283
Db	1030	CTTGATGCTGCCATTTGGGTATAGAAATTTGGAGGAAGTTTGTAT	1074
RESULT	16		
CRUS0832			
LOCUS		1278 bp DNA linear BCT 14-JUL-1996	
DEFINITION		Cowdria ruminantium Gardel major antigenic protein 1 (map1) gene, complete cds.	
ACCESSION		U50832	
VERSION		U50832.1	GI:1418253
KEYWORDS			
SOURCE		heartwater rickettsia strain-Gardel.	
ORGANISM		Cowdria ruminantium	
REFERENCE		Bacteria; Proteobacteria; alpha subdivision; Rickettsiales; Rickettsiaceae; Ehrlichiae; Cowdria.	
AUTHORS		1 (bases 1 to 1278) Reddy,G.R., Sulsona,C.R., Harrison,R.H., Mahan,S.M., Burridge,M.J. and Barbet,A.F.	
TITLE		Sequence heterogeneity of the major antigenic protein 1 genes from Cowdria ruminantium isolates from different geographical areas Clin. Diagn. Lab. Immunol. 3 (4), 417-422 (1996)	
JOURNAL			
MEDLINE		96400830	


```
Db 322 GCTGACGAATTGGATACTCAATGAAT---GGGCCAAGAATAGAGTTTGAAGTATCCTAT 378
Qy 119 GluThrPheHisIleLysAsnGly-----TyrLysArgIleAspCysGluLysHis 136
Db 379 GAAACTTTTGATGTCAAAATCCAGGTGGCACTATAAATAAT-----GATGCCAC 429
Qy 137 PheAlaLeuAlaLysGluIleSerGlyCysSerGlyCysProAlaAsnLysTyrVal 156
Db 430 ATGTACTGTGCTTTAGACAGCA-----CAGCCAACAGTAGTAATCAAGGGCA 477
Qy 157 ThrLeuIle-----AsnAsnGlyIleSerLeuThrSerAlaLeuIle 170
Db 478 ACATTAGCTTCATCTGTTATGGTAAAAACGAAATTTAACAGATATATCAATATGTTA 537
Qy 171 AsnValCysTyrAspVal-----AspGlyLeuLysHisAsnIleIleThrTyrSerCys 188
Db 538 AATCATGTTATGATATAATAGCTTGTGGAATGCCA-----GTTCTCCATATGTATGT 591
Qy 189 LeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyr 208
Db 592 GCAGTATTGGTACTGATTTAGTATCATAGTAATTAATGCTACAAATCCTAAATATCTAT 651
Qy 209 GlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerValPheIleGluGly 228
Db 652 CAAGGAAAGCTAGGCATAAGTTACTCAATCAATCCTGAAGCTTCTATCTTTATTGGTGG 711
Qy 229 TyrTyrHisGlyLeuPheGlyLysPheGluLysIleProValAsnTyrProCysAsp 248
Db 712 CATTTCCATAGATATAGGTAATGAATTAAGATATGCTACTCAAAA-----ATA 765
Qy 249 TyrProSerProThrProAsnSerLysProHisValHisThrThrAlaLeuAlaMet 268
Db 766 TTTACTTCAACTAGTAAGCATCATCTATACCTAAT---CCTGGCTTTGCATCAGCAACA 822
Qy 269 LeuSerIleGlyTyrTyrGlySerIleGlyLysPheIle 283
Db 823 CTTGATGCTGCCATTCGGTATAGAAATTGGAGGAAGGTTTGTGA 867

RESULT 19
AF368007 828 bp DNA linear BCT 06-NOV-2001
LOCUS Cowdria ruminantium isolate Mali major antigenic protein MAP1 gene,
DEFINITION partial cds.
ACCESSION AF368007
VERSION AF368007.1 GI:15429325
KEYWORDS heartwater rickettsia.
SOURCE Cowdria ruminantium
ORGANISM Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
REFERENCE 1 (bases 1 to 828)
AUTHORS Allsopp,M.T.E.P., Dorfling,C.M., Maillard,J.C., Bensaid,A.,
Haydon,D.T., van Heerden,H. and Allsopp,B.A.
TITLE Ehrlichia ruminantium Major Antigenic Protein Gene (map1) Variants
Are Not Geographically Constrained and Show No Evidence of Having
Evolved under Positive Selection Pressure
J. Clin. Microbiol. 39 (11), 4200-4203 (2001)
JOURNAL 21539003
MEDLINE 11682561
PUBMED
REFERENCE 2 (bases 1 to 828)
AUTHORS Allsopp,M.T., Dorfling,C., Maillard,J.-C., Bensaid,A., Van
Heerden,H. and Allsopp,B.A.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2001) Molecular Biology, Onderstepoort Veterinary
Institute, Old Soutpansberg Road, Onderstepoort, Gauteng 0110,
South Africa
FEATURES
source Location/Qualifiers
1. 828
/organism="Cowdria ruminantium"
/isolate="Mali"
/db_xref="taxon:779"
<1..>828
/codon_start=1
CDS
```

```
/transl_table=11
/product="major antigenic protein MAP1"
/protein_id="AAK98147.1"
/db_xref="GI:15429326"
/translation="FISTLILSVFLPGVFSQVSDVIOEDNSPVGSVVISAKYMTASH
FGMSIKEDSKNTQTVFGLKKDWDGVKTPSDSNHSTIFTEKDYISRYENNPFLGFAG
AICYSNMGPRIEFEVYETPDVKPNPGNYKNDAMYCALDTAAQNSTNGAGLTTSMV
KNLNTDLSLMLNACYDILMDGMPVSPVCAGICTDLVSINATNPKLSYQGLGISY
SINPEASIFIGGHFHRVIGNEFKDIATSKTFTSGTGISIPNPGFASALDVCHEGIEV
"
BASE COUNT 292 a 141 c 142 g 253 t
ORIGIN
Alignment Scores:
Pred. No.: 1,92e-26 Length: 828
Score: 389.00 Matches: 91
Percent Similarity: 54.07% Conservative: 42
Best Local Similarity: 36.99% Mismatches: 87
Query Match: 25.76% Indels: 26
DB: 1 Gaps: 10
US-09-846-808-1 (1-284) x AF368007 (1-828)
Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
Db 19 ATATCATTAGTGTCTTCTTACCTGGTGTCTTTTCTGATGTAATACAAAGAATAAC 78
Qy 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
Db 79 AGCCAGTAGTAGTCGTTTACATAGTGCAAAGATACATGCCCACTGCTTCACACTT 138
Qy 49 LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysAsp 68
Db 139 AAAATGTCATAAAGGAAGATTCAAGAAATFACCTCAACACAGTATTTGGCTTAAAAAAGAT 198
Qy 69 LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsnPheAsnIle 85
Db 199 -----TGGGATGGGTTAAAAACACCATCAGATCTTAGCAATCATCTCTACA 243
Qy 86 -----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPheSer 99
Db 244 ATTTTCTAGTAAAAAGATATCTTCAATATGAAACAAATCCCTTTTATAGGTTTCGCT 303
Qy 100 GlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGlu 119
Db 304 GGACCAATTGGATACTCAATGAAT---GGACCAAGAATAGAGTTTGAAGTATCCTATGAA 360
Qy 120 ThrPheHisIleLysAsnGly-----TyrLysArgIleAspCysGluLysHisPhe 137
Db 361 ACTTTTGTATGTCAAAAATCCAGCGGTAATTTATAAAAC---GATGCCACAGATCTACTGT 417
Qy 138 AlaLeuAlaLysGluIleSerGlyCysSerAsnAsnProAlaAsnLysTyrValThr 157
Db 418 GCTTTAGATACAGCAGCAACAAACAGCACTAATGGTCAGGATTAACATCATCTGTTATG 477
Qy 158 LeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal--- 176
Db 478 GTAAAAACGAAATTTAACAGATATATCACTAATGTTAAATGCGTATGATATAATG 537
Qy 177 ---AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThr 195
Db 538 CTTGATGGAATGCCA-----GTCTCCCATATGTATGTCAGGATTTTGGCAGCATGATCTA 591
Qy 196 IleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSer 215
Db 592 GTGCAGTAATTAATGCTACAAATCCTAAACTATCTTACCAGGCACTTAGGTATTAGC 651
Qy 216 TyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGly 235
Db 652 TATTCAATTAATCCAGAGCATCTATATTTATTGGAGGACACTTTTCATAGAGTTATAGGT 711
Qy 236 LysLysPheGluLysIle 241
:::|||||::: |||
```



```

/codon_start=1
/transl_table=11
/product="major antigenic protein 1"
/protein_id="AAC44146.1"
/db_xref="GI:1418260"
/transl="MNCKRIFITSTLISLVSFLPGVSFSDVIOEDSNPAGSVVISAKY
MPTASHFGKMSIKEDSRDTPKAVEGLKDKWDGVKTPSSDGSNSIIFTEKDYSPKYENN
PLGAGAVGYSNMGPRIEFVSYETFDVKNPGGYNKNDAMHVCALDTCGPSTOGAT
LNSVMVKNENLTDIALMNCYDITLECMPPSPYVCAGIGTDLSVINATNPKLSYQ
GKLGISYNPEASIFIGGHRVHVGNEFKDITTSKIFTSTGKLATAASPGFASATLD
VCHGIEIGGRFVF"
BASE COUNT 431 a 175 c 207 g 465 t
ORIGIN

Alignment Scores:
Pred. No.: 4,01e-26 Length: 1278
Score: 388.00 Matches: 97
Percent Similarity: 51.93% Conservative: 51
Best Local Similarity: 34.04% Mismatches: 115
Query Match: 25.70% Indels: 22
DB: 1 Gaps: 10

US-09-846-808-1 (1-284) x CRU50835 (1-1278)
QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeu----- 28
Db 432 ATATCATTAGTGTCAATTTTACCTGGTGTCTCTTTCTCATGTATATACAGGAGTAGC 491
QY 29 AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
Db 492 AACCCACGGGTAGTGTTCATATGCGCAAAATACATGCACTGCTTCACATTTTGGT 551
QY 49 LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68
Db 552 AATGTCTCATCAAGAGAGATTCAAAAATACCTCAACACGATTTGGCCTAAAAAAGAT 611
QY 69 LeuLeuAsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle----- 85
Db 612 ---TGGATGGCGTTAAACACACCACCATCATCAGATGCGGTAGCAATAGTATTATCTTCACT 668
QY 86 -----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePhe 102
Db 669 GAAAAGACTATTCTCAATTAATGAACAACACCCCATTTTATGGTTTGTCTGGAGCAATT 728
QY 103 GlyTyrTyrAsnLysAsnPheArgIleGluSerGlyLeuSerTyrGluThrPheHis 122
Db 729 GGGTATTCAATGAAT---GGTCCAGAAGATAGAAATTTGAAGTCTCTATGAACCTTTTGT 785
QY 123 IleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHisPheAlaLeuAla 140
Db 786 GTAAAAAACCCAGGTGGTGAATTTATAAAAT---GATGCACATATGATTTGCTCTAGAT 842
QY 141 LysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsn 160
Db 843 ACAGGAACACCGAGATCTACTAGGGTGCAACATTAATTCATCTGTTATGGTAAATAAAT 902
QY 161 AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCystTyrAspVal-----AspGly 178
Db 903 GAAATTTAACTGATATGCACTAAATGTTAAATGCATGCTATGATATACACTTGAAGGA 962
QY 179 LeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPhe 198
Db 963 ATGCCA-----GTTTCTCCATATGTTGTGTCAGGCATTTGGTACTGATTTAGTGTAGTA 1016
QY 199 LeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrVal 218
Db 1017 ATCAATGCTACAATCTCAAGTTATCTTATCAGGGAAGTTAGGTATTTAGTTATCAATA 1076
QY 219 SerProGlnValSerValPheIleGluGlyTyrThrGlyLeuPheGlyLysLysPhe 238
Db 1077 AATCCCTCAAGCTTCTATCTTTATTTGGTGGACATTTCCATAGAGTTATAGTAAATGATTT 1136
QY 239 GluLysIleProValAsnTyrProCysAspTyrProSerProThrProAsnSerLys 258

```

```

Db 1137 AAACACATTACTACTTCCAAAATATTACCTCAACTGGTAAATTAGCTACTGCAGCTAGC 1196
QY 259 ProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIle 278
Db 1197 CCAGGTTT-----GCATCAGCAACACTTGATGTTTGCCATTTCCGTATAGAAATT 1247
QY 279 GlyIleLysPheIle 283
Db 1248 GGAGGAAGGTTTGTA 1262

RESULT 22
CRMAP1
LOCUS CRMAP1 1467 bp DNA linear BCT 10-OCT-1994
DEFINITION C.ruminantium map1 gene.
ACCESSION X74250
VERSION X74250.1 GI:454266
KEYWORDS immunodominant protein; major antigen; map1 gene; outer membrane protein.
SOURCE heartwater rickettsia.
ORGANISM Cowdria ruminantium
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales; Rickettsiaceae; Ehrlichiae; Cowdria.
REFERENCE 1 (bases 1 to 1467)
AUTHORS Van Vliet,A.H.M.
TITLE Direct Submission
JOURNAL Submitted (22-JUL-1993) A.H.M. Van Vliet, Inst of Infectious Diseases & Immunology, Dept of Bacteriology, School of Veterinary Medicine, Yalelaan 1, PO Box 80.165, 3508 TD Utrecht, NETHERLANDS
REFERENCE 2 (bases 1 to 1467)
AUTHORS Van Vliet,A.H., Jongejan,F., van Kleef,M. and van der Zeijst,B.A.
TITLE Molecular cloning, sequence analysis, and expression of the gene encoding the immunodominant 32-kilodalton protein of Cowdria ruminantium
JOURNAL Infect. Immun. 62 (4), 1451-1456 (1994)
MEDLINE 94178956
FEATURES
source
1..1467
/organism="Cowdria ruminantium"
/isolates="Senegal stock"
/db_xref="taxon:779"
481..1335
/genes="map1"
481..1335
/genes="map1"
/codon_start=1
/transl_table=11
/product="major antigenic protein"
/protein_id="CAA52309.1"
/db_xref="GI:454267"
/translation="MNCKRIFITSTLISLVSFLPGVSFSDVIOENNPVGSVVISAKY
MPTASHFGKMSIKEDSRDTPKAVEGLKDKWDGVKTPSGNTNSIIFTEKDYSPKYENNPL
GFAGAVGYSNMGPRIEFVSYETFDVKNPGGYNKNDAMHVCALDTCASSYAGATNTSYM
VKNENLTDISLMACDIDMLDMPVSPYVCAGIGTDLSVINATNPKLSYQOKLGIS
YSINPEASIFIGGHRVHVGNEFKDITTSKIFTSSGNASSAVSPGFASALDVCHEFI
EIGGRFVF"
BASE COUNT 493 a 204 c 242 g 528 t
ORIGIN

Alignment Scores:
Pred. No.: 4,74e-26 Length: 1467
Score: 388.00 Matches: 100
Percent Similarity: 52.10% Conservative: 49
Best Local Similarity: 34.97% Mismatches: 107
Query Match: 25.70% Indels: 30
DB: 1 Gaps: 13

US-09-846-808-1 (1-284) x CRMAP1 (1-1467)
QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn 32
Db 517 ATATCATTAGTGTCAATTTTACCTGGTGTGTCATTTTCTGATGTAATACAGGAAGAAC 576

```

```
QY 33 -----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
Db 577 AATCCAGTAGTAGTGTATATATAGCCCAAAATACATGCCCACTGCATCACATTTTGGT 636
QY 49 LysLeuLeuIleGlySerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68
Db 637 AAAATGTCATTAAGACAGATCTCTAGAGATACATAAGCAGATTTTGGCTTAAAGAAAT 696
QY 69 LeuLeuAsnAspLeuThrGlyIleLys-----AspAsnThrAsn-----Phe 83
Db 697 -----TGGGATGGAGTTAAACACCACTCGGTAACACCAATTCATTTT 741
QY 84 AsnIleLys---TyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIle 101
Db 742 ACTGAAAAAGACATATCTTCAATTAAGAAATACCAATTTTATAGTTTTCAGGAGCA 801
QY 102 PheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThrPhe 121
Db 802 GTTGATACTCAATGAAT---GGACCAAGATAGAAATTTGAAGTATCTTATGAAACTTTC 858
QY 122 HisIleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHisPheAlaLeu 139
Db 859 GAGCTAAGAAATCCAGCGGTAACTACAAAAC-----GATGCACATATGTTTCT 909
QY 140 AlaLysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuIle 159
Db 910 GCTCTAGATACAGCAAGCTCTACTGCAGGAGCAACTACATCT---GTTATGTGTA 966
QY 160 AsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal-----Asp 177
Db 967 AATGAAATTTTAAACAGATATTTTCAATTAATGCTTAAACGCGATGTTATGACATA 1026
QY 178 GlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAsp 197
Db 1027 GGAATGCCA-----GTTTCTCCATATGATGTCGAGCAATGGTACTGACATTAGTATCA 1080
QY 198 PheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThr 217
Db 1081 GTAATTAAGCGTACAAATCTTAATTTATCAAGGAAATTAGGAATAAGTTATTATCA 1140
QY 218 ValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLys 237
Db 1141 ATAAACCCGGAAGCTCTATCTTTTATTTGGTGGCACTTCCATAGAGTATCAGTAA 1200
QY 238 PheGluLysIleProValAsnTyrProCysAspTyrProSerProThrProAsnSer 257
Db 1201 TTTAAGATATTCGCTACTCTAAAGTTTCTACTAGCAGTGGTAATGCCAGTAGTCTGTT 1260
QY 258 LysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlySer 277
Db 1261 AGTCCAGGTTTT-----GCATCAGCAATACTTGTATGTTTGGCACTTCGGCATAGAA 1311
QY 278 IleGlyIleLysPheIle 283
Db 1312 ATTGGAGGAAGGTTTGTA 1329
RESULT 23
AX042313 AX042313 837 bp DNA linear PAT 23-NOV-2000
LOCUS Sequence 9 from Patent WO0065063.
DEFINITION AX042313
ACCESSION AX042313
VERSION AX042313.1 GI:11340973
KEYWORDS
SOURCE
ORGANISM
Ehrlichia chaffeensis.
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
REFERENCE
1 (bases 1 to 837)
AUTHORS
Barbet A.F., Bowie M.V., Ganta, R.R., Burrige M.J., Mahan, S.M.,
Mcquire, T.C., Rurangirwa, F.R., Moreland, A.L., Simbi, B.H.,
Whitmire, W.W. and Allenman, A.R.
TITLE
Nucleic acid vaccines against rickettsial diseases and methods of
```

```
use
JOURNAL Patent: WO 0065063-A 9 02-NOV-2000;
UNIVERSITY OF FLORIDA (US)
FEATURES Location/Qualifiers
source 1..837
/organism="Ehrlichia chaffeensis"
/db_xref="taxon:945"
BASE COUNT 279 a 130 c 159 g 269 t
ORIGIN
```

```
Alignment Scores:
Pred. No.: 2,66e-26 Length: 837
Score: 387.50 Matches: 102
Percent Similarity: 51.05% Conservative: 44
Best Local Similarity: 35.66% Mismatches: 117
Query Match: 25.66% Indels: 23
DB: 6 Gaps: 10
```

US-09-846-808-1 (1-284) x AX042313 (1-837)

```
QY 7 ArgLysPheValLeu-----TrpValMetLeuIleLeuPheThrProHisIleSer 23
Db 12 AAAAAATTTTATAACAACATGCATTAGTATCACTAATGCTCTTCTACTCGGAATATCA 71
QY 24 LeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrValGlyIleGlnTyr 40
Db 72 TTTTCTGATCCAGTCAAGTGCACAATATAGTGGTAAATTTCTATGTTAGTGCAAGTAT 131
QY 41 LysProAlaArgGlnHisLeuSerLysLeuIleLysGluSerAlaAlaAsnThrVal 60
Db 132 ATGCCAAGTGTCTGCGATTTTGGCATTTTCTGCCAAGAAAGAAAAATCCTACTGTT 191
QY 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuThrGlyIleLysAspAsn 80
Db 192 GCATTGTATGGCTTAAACAAGATTGGGAGGAGTTAGCTCATCAAGTCACAATGATAAT 251
QY 81 ThrAsnPheAsnIleLys---TyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
Db 252 ---CATTTCAATAACAAGGTTTATTCATTTAAATATGAAATAAACCCATTTTATAGGTTT 308
QY 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleAspCysGluLysSerTyr 118
Db 309 CGAGGAGCTATTGGTTATTCAATGGGT---GGTCCAAGAGTAGAGTTTGAAGTGTCTAT 365
QY 119 GluThrPheHisIleLysAsn-----AsnGlyTyrLysArgIleAspCysGluLysHis 136
Db 366 GAAACATTTGAGCTTAAATAATCAGGGTAACTATAATAAAT---GATGCTCACAGATAC 422
QY 137 PheAlaLeuAlaLysGluIleSerGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
Db 423 TGTGCTTTAGGTCAACAGACACACAGCGGAATA-----CCTAAATACTAGTAATACGTA 476
QY 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
Db 477 CTGTTAAAAAGCGAAGGATTGCTTGACATATCATTTATGCTAAATGCATGCTATGATA 536
QY 177 AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAspThrIle 196
Db 537 ATAACAGAGAGCATACCTTTGCTCTCTTACATATGTCAGAGTCTTGGTACTGATTAA 596
QY 197 AspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyr 216
Db 597 TCCATGTTTGAAGCTACAAATCTTAAATTTTCTACCAAGGGAAGTTAGGTCTTAAGTTAC 656
QY 217 ThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLys 236
Db 657 TCTATAAACCCAGAGAGCTTCTGTTATTATTGTTGGACATTTTCATAGGTGATAGGAAC 716
QY 237 LysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrProAsn 256
Db 717 GAATTTAGGACATCTCTACTCTGAAGAGCATTTGTTTACGTCATCAGCTACTCCA----- 770
QY 257 SerLysProHisValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGly 276
```


Branch, 301 Univ. Blvd., Galveston, TX 77555-0609, USA

FEATURES

Source
1. .1307
/organism="Ehrlichia chaffeensis"
/strain="V8"
/db_xref="taxon:945"
236. .1078
/note="p28-19; 28 kDa outer membrane protein"
/codon_start=1
/transl_table=11
/product="outer membrane protein p28"
/protein_id="AAL12924.1"
/db_xref="GI:15991542"
translation="MNKKVFITSLISLPGVFSFSDPTGSGINGNFYISGKYP
SASHFGVSAKEERTTAGVFLGDKWDGSAISHPTPENITVSNYSKYENNPFLGF
AGAIGYMDGPRIELEVSEYFDVKNQNSYKNEAHRYCALSHHSSGTSMSAGDKFV
FLKNGLLDVSPMLNACYDVLSEGIPIFPSPYICVIGTDLVSMFEATNPKISVQGLGL
SYSISPETSVFTGGHFKHVGNEFRDIPAVIPSGSSLTGNHPAIVLVDVCHFGIELGG
RFAF"

BASE COUNT 420 a 212 c 245 g 430 t

ORIGIN

Alignment Scores:
Pred. No.: 6,25e-26 Length: 1307
Score: 386.00 Matches: 95
Percent Similarity: 51.44% Conservative: 48
Best Local Similarity: 34.17% Mismatches: 115
Query Match: 25.56% Indels: 20
DB: 1 Gaps: 9

US-09-846-808-1 (1-284) x AF393394 (1-1307)

QY 13 ValMetLeuIleuPheThrProHisIleSerLeuAlaSerValLeuAsnAsp----- 30
Db 272 ATATCACTAATCTCTCTACCTGGAGTATCATTTCTGACCCACACAGTAGTGAT 331
QY 31 HisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeu 50
Db 332 AACGGCAATTTATATCATGTCAGTGGAATAATATATGCCAAGCGCTTCGCATTTGGGGTGT 391
QY 51 LeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeu 70
Db 392 TGTGCTAAGGAAGAAAGAGTACACAGCTGGAGTATTTGGATTGAAGCAAGATTGGGAT 451
QY 71 AsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsnPro 89
Db 452 GGCAGTGCAATATCTCACACCACCCAGAAATAATATATTCAGTCTTCAATATTCTGTT 511
QY 90 TyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsnLys 108
Db 512 AAATATGAAATAAACCCATTTTAGGATTTCGAGGAGCTATTGTTACTCAATGGAT--- 568
QY 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn-----Asn 126
Db 569 GGGCCCAAGAAATAGAGCTTGAAGTATCTTATGACACATTCGATGTAAATAATCAAGTAA 628
QY 127 GlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGly 146
Db 629 AGTTACAAGAAC---GAACGGCATAGGTATTGTGTTTATCATCATCAGATTCAAGAA 685
QY 147 SerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThr 166
Db 686 AGCATCTCAGTGCAGGTATAAATTTGTTTCTAAAAAATGAAGGATTACTTACGTA 745
QY 167 SerAlaLeuIleAsnValCysTyrAspVal-----AspGlyLeuLysHisAsnIleIle 184
Db 746 TCGTTTCATGCTGAACCGCATGCTATGATGTAATAAGTGAAGGAATACCTTTTCT----- 799
QY 185 ThrTyrSerCysLeuGlyPheGlyValAlaAspThrIleAspPheLeuSerLysTyrThr 204
Db 800 CCTTACATATGTAGTATCGGAGCTGATTTAGTATCCATGTTTGAAGCTACAAATCCT 859
QY 205 LysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerVal 224

Db 860 AAAATTCTTACCAAGAAAGTTAGTTTAAAGCTACTATATAAGCCAGAAACTTCTGTG 919
QY 225 PheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLysIleProValAsn 244
Db 920 TTTATTGGTGACATTTTTCATAAGGTGATAGGAACAATTTAGA----- 964
QY 245 TyrProCysAspTyrProSerProThrProProAsnSerLysProHisValHisThrThr 264
Db 965 -----GATATTCTCTGCTGTAATACCTAGTGGATCAAGTCTTACAGGAACCAACCC 1015
QY 265 AlaLeuAlaMetLeuSerIleGlyTyrTyrGlySerIleGlyLysIleLysPhe 282
Db 1016 GCAATAGTAACACTGGACGTATGCCACTTTGTTGATAGAGTTGGAGGAAGTTT 1069
RESULT 26
AF393390
LOCUS
DEFINITION
AF393390 . 1309 bp DNA linear BCT 10-OCT-2001
Ehrlichia chaffeensis strain V4 outer membrane protein p28 gene,
complete cds.
ACCESSION AF393390 GI:15991533
KEYWORDS
SOURCE
ORGANISM
Ehrlichia chaffeensis.
Ehrlichia chaffeensis
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
REFERENCE
1 (bases 1 to 1309)
AUTHORS
Long,S.W., Zhang,X.-F., Qi,H., Standaert,S., Walker,D.H. and
Yu,X.-J.
TITLE
Allele variation and patterns of transcription of the Ehrlichia
chaffeensis 28 kDa outer membrane protein multigene family
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1309)
AUTHORS
Yu,X.-J., Zhang,X.-F. and Walker,D.H.
TITLE
Direct Submission
JOURNAL
Submitted (20-JUN-2001) Pathology, University of Texas Medical
Branch, 301 Univ. Blvd., Galveston, TX 77555-0609, USA
FEATURES
Location/Qualifiers
1. .1309
/organism="Ehrlichia chaffeensis"
/strain="V4"
/db_xref="taxon:945"
238. .1080
/note="p28-19; 28 kDa outer membrane protein"
/codon_start=1
/transl_table=11
/product="outer membrane protein p28"
/protein_id="AAL12920.1"
/db_xref="GI:15991534"
/translation="MNKKVFITSLISLPGVFSFSDPTGSGINGNFYISGKYP
SASHFGVSAKEERTTAGVFLGDKWDGSAISHPTPENITVSNYSKYENNPFLGF
AGAIGYMDGPRIELEVSEYFDVKNQNSYKNEAHRYCALSHHSSGTSMSAGDKFV
FLKNGLLDVSPMLNACYDVLSEGIPIFPSPYICVIGTDLVSMFEATNPKISVQGLGL
SYSISPETSVFTGGHFKHVGNEFRDIPAVIPSGSSLTGNHPAIVLVDVCHFGIELGG
RFAF"

BASE COUNT 420 a 212 c 247 g 430 t

ORIGIN

Alignment Scores:
Pred. No.: 6,27e-26 Length: 1309
Score: 386.00 Matches: 95
Percent Similarity: 51.44% Conservative: 48
Best Local Similarity: 34.17% Mismatches: 115
Query Match: 25.56% Indels: 20
DB: 1 Gaps: 9

US-09-846-808-1 (1-284) x AF393390 (1-1309)

QY 13 ValMetLeuIleuPheThrProHisIleSerLeuAlaSerValLeuAsnAsp----- 30
Db 274 ATATCACTAATCTCTCTACCTGGAGTATCATTTTCTGACCCACACAGTAGTGAT 333

Qy 31 HisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeu 50
Db 334 AACGCCAATTTCTATATCAGTGGAAATATATATGCAAGCGCTTGCATTTGGGGTGTTT 393
Qy 51 LeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeu 70
Db 394 TCTGCTAAGGAAGAAAGATGACACAGCTGGAGTATTGGATTGAAGCAAGATTGGGAT 453
Qy 71 AsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsnPro 89
Db 454 GGCAGTGCATATCTCACACACCCCAAAATATATTCACCTGTTCAAAATATTCGTTT 513
Qy 90 TyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsnLys 108
Db 514 AAATATGAAATAACCCATTTTGGAGTATTCGAGGAGCTATTGGTTACTCAATGAT--- 570
Qy 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn-----Asn 126
Db 571 GGCCCAAGAATAGAGCTTGAAGTATCTTATGAGACATTCGATGTAAATAAATCAAGGTAA 630
Qy 127 GlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGly 146
Db 631 AGTTACAAGAAC---GAAGCCATAGGTATTGCTTTATCTCATCAGTTCAGGAACA 687
Qy 147 SerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThr 166
Db 688 AGCATGTCGAGTCGAGGTGATAAATTTGTTTCTTCTAAAAAATGAAGGATTACTGCACGTA 747
Qy 167 SerAlaLeuIleAsnValCysTyrAspVal-----AspGlyLeuLysHisAsnIleIle 184
Db 748 TCGTTCATGCTGAACGATGCTATGATGATAAAGTGAAGAAATACCTTTTCT----- 801
Qy 185 ThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSerLysTyrThrThr 204
Db 802 CCTACATATGCTAGGTATCGGACTGATTTAGTATCCAGTTTGAAGCTACAAATCCT 861
Qy 205 LysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerProGlnValSerVal 224
Db 862 AAAATTTCTTACCAAGAAAGATTAGGTTTAAGCTACTATATAAGCCCAAGAAATCTCTG 921
Qy 225 PheIleGluGlyTyrTyrHisGlyLysPheGlyLysLysPheGluLysIleProValAsn 244
Db 922 TTTATTTGTGTGACATTTTTCATAAGGTGATAGGAACGAATTTTGA----- 966
Qy 245 TyrProCysAspTyrProSerProThrProProAsnSerLysProHisValHisThrThr 264
Db 967 -----GATATTCCTGCTGATATACCTAGTGGATCAAGTCTTACAGGAACACCCCT 1017
Qy 265 AlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyIleLysPhe 282
Db 1018 GCAATAGTAACACTGGAGGTATGCCACTTTGGTATAGACTTGGAGGAAGGTTT 1071
RESULT 27
AF077733
LOCUS AF077733 1243 bp DNA linear BCT 20-JUN-2001
DEFINITION Ehrlichia chaffeensis strain Jax outer membrane protein p28
ACCESSION AF077733
VERSION AF077733.2 GI:14495545
KEYWORDS Ehrlichia chaffeensis.
SOURCE Ehrlichia chaffeensis
ORGANISM Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
REFERENCE 1 (bases 1 to 1243)
AUTHORS Yu,X.-J., McBride,J.W. and Walker,D.H.
TITLE Genetic diversity of the 28-kilodalton outer membrane protein gene in human isolates of Ehrlichia chaffeensis
J. Clin. Microbiol. 37 (4), 1137-1143 (1999)
JOURNAL 99175287
MEDLINE 10074538
PUBMED
REFERENCE 2 (bases 1 to 1243)
AUTHORS Yu,X.-J. and Walker,D.H.

TITLE Direct Submission
JOURNAL Submitted (14-JUL-1998) Pathology, The University of Texas Medical Branch, 301 Univ. Blvd., Galveston, TX 77555, USA
REFERENCE 3 (bases 1 to 1243)
AUTHORS Yu,X.-J. and Walker,D.H.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2001) Pathology, The University of Texas Medical Branch, 301 Univ. Blvd., Galveston, TX 77555, USA
REMARK Sequence update by submitter
COMMENT On Jun 20, 2001 this sequence version replaced gi:3414968.
FEATURES
source
1..1243
/organism="Ehrlichia chaffeensis"
/strain="Jax"
/db_xref="taxon:945"
237..1079
/codon_start=1
/transl_table=11
/product="outer membrane protein p28 precursor"
/protein_id="AAC31546.1"
/db_xref="GI:3414969"
/translation="MNVKKVFITTSALISLISLPQVSFSDPTGSGINGNFVIGSKYMP
SASHFGVPSAKEERSTTAGVGLQDWDGSAISHTTPTNIFTVSNYSFKYNNPPLCF
AGAIGSMDGPRIELEVSYETFDVKNQNSYKNAHRYCALSHSSGSSMSAGDKFY
FLKNEGLLDSFMLNACYDVISGPIYICVIGITDLVSFMFATNPNTKISYQKGL
SYSTSPETSVFIGGFHKVIGNEFRDIPAVIPNGSSLTGNHPAIVTLDVCHFGIELGG
RFAP"
BASE COUNT 396 a 197 c 233 g 417 t
ORIGIN
Alignment Scores:
Pred. No.: 7.25e-26 Length: 1243
Score: 385.00 Matches: 95
Percent Similarity: 51.44% Conservative: 48
Best Local Similarity: 34.17% Mismatches: 115
Query Match: 25.50% Indels: 20
DB: 1 Gaps: 9
US-09-846-808-1 (1-284) x AF077733 (1-1243)
Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAsp----- 30
Db 273 ATATCACAATATCTCTCTACCTGGAGTATCATTTCTGACCCCAACAGGTAGTGGTATT 332
Qy 31 HisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeu 50
Db 333 AACGCCAATTTCTATATCAGTGGAAATATATATGCAAGCGCTTGCATTTGGGGTGTTT 392
Qy 51 LeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAspLeuLeu 70
Db 393 TCTGCTAAGGAAGAAAGATGACACAGCTGGAGTATTGGATTGAAGCAAGATTGGGAT 452
Qy 71 AsnAspLeuLeuThrGlyIleLysAspAsnThrAsnPheAsnIle---LysTyrAsnPro 89
Db 453 GGCAGTGCATATCTCACACACCCCAAAATATATATTCACCTGTTTCAAAATATTCGTTT 512
Qy 90 TyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsnLys 108
Db 513 AAATATGAAATAACCCATTTTGGAGTATTCGAGGAGCTATTGGTTACTCAATGAT--- 569
Qy 109 AsnPheArgIleGluSerGluLeuSerTyrGluThrPheHisIleLysAsn-----Asn 126
Db 570 GGCCCAAGAATAGAGCTTGAAGTATCTTATGAGACATTCGATGTAAATAAATCAAGGTAA 629
Qy 127 GlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAlaLysGluIleSerGlyGly 146
Db 630 AGTTACAAGAAC---GAAGCCATAGGTATTGCTGCTTTATCTCATCAGATTCAGGAACA 686
Qy 147 SerAsnAsnProAlaAsnAsnLysTyrValThrLeuIleAsnAsnGlyIleSerLeuThr 166
Db 687 AGCATGTCGAGTCGAGGTGATAAATTTGTTTCTTCTAAAAAATGAAGGATTACTGCACGTA 746
Qy 167 SerAlaLeuIleAsnValCysTyrAspVal-----AspGlyLeuLysHisAsnIleIle 184

TITLE	Genetic diversity of the 28-kilodalton outer membrane protein gene in human isolates of <i>Ehrlichia chaffeensis</i>
JOURNAL	J. Clin. Microbiol. 37 (4), 1137-1143 (1999)
MEDLINE	99175287
PUBMED	10074538
REFERENCE	2 (bases 1 to 2977)
AUTHORS	Yu, X. J., McBride, J. W., Zhang, X. and Walker, D. H.
TITLE	Characterization of the complete transcriptionally active <i>Ehrlichia chaffeensis</i> 28 kDa outer membrane protein multigene family
JOURNAL	Gene 248 (1-2), 29-68 (2000)
REFERENCE	3 (bases 1 to 1268)
AUTHORS	Yu, X. J., McBride, J. W. and Walker, D. H.
TITLE	Direct Submission
JOURNAL	Submitted (26-MAY-1998) Pathology, The University of Texas Medical Branch, 301 Univ. Blvd., Galveston, TX 77555, USA
REFERENCE	4 (bases 1 to 2977)
AUTHORS	Yu, X. J., McBride, J. W., Zhang, X. and Walker, D. H.
TITLE	Direct Submission
JOURNAL	Submitted (01-FEB-2000) Pathology, University of Texas Medical Branch, 301 Univ. Blvd., Galveston, TX 77555, USA
REMARK	Sequence update by submitter
COMMENT	On May 24, 2000 this sequence version replaced gi:3192922.
FEATURES	Location/Qualifiers
source	1. .2977 /organism="Ehrlichia chaffeensis" /strain="Arkansas" /db_xref="taxon:945" /complement(180. .995) /gene="p28-20" /complement(180. .995) /genes="p28-20" /notes="component of 28-kDa outer membrane protein multigene family" /codon_start=1 /transl_table=11 /product="p28-20 precursor" /protein_id="AAC19134.2" /db_xref="GI:8052525" /translation="MNVKFFVGVGVALTLLSFLPDNSFSDANVPEGRKGFYVGTQYKV GVDFNFSAEETPLGLTKSIFALGDKSIDSHAGFTQAYNPTYASNFAEGGVIGY YVNDVFEPEGAYENEPERQWYPEGESHKFPALSRSTVQDNKEFIVLENDGVIDKS LVNVCIDTAHGSIPAPACMGAGADIYKIGISLPKFSYQVKFGVNIPTVSVNVMFL GGGYHKYIGNRYVEVYIAHPATLNVPKTTSASATLDTDPFGEVGMRFLL" /complement(183. .920) /genes="p28-20" /product="p28-20" /complement(921. .995) /genes="p28-20" /complement(1178. .2014) /gene="p28-21" /complement(1178. .2014) /genes="p28-21" /notes="component of 28-kDa outer membrane protein multigene family" /codon_start=1 /transl_table=11 /product="p28-21" /protein_id="AAF71791.1" /db_xref="GI:8052526" /translation="MYRISVITLMLLVPCDCSGSLDVSLSNRLKPVFLGIGYK LSAPLFSFSIGTYRINGVKTDRVVGKLSIDLDADKAMKDFNNFNSEEVYPKVDY NIETGLSFYGSFNRILVELEGSKKFDVLDTRNHLVNDYRHIALVRSNPPTLYDF VLKNDGVEYFTLLINCYFDVDTNIPVFCVIGEDIIKIFDSIRKFSFNKLGKIG YLMNQDMLFFDDVYHVRVGNVYNNIPVQVSLPNLSTAAKLDMEYFGAIGIKV FV"
BASE COUNT	1172 a 529 c 353 g 923 t
ORIGIN	
Alignment Scores:	2.58e-25 Length: 2977
Pred. No.:	384.00 Matches: 95
Score:	50.00% Conservative: 47
Percent Similarity:	33.45% Mismatches: 106
Best Local Similarity:	


```
/protein_id="CAC17072.1"
/db_xref="GI:11340966"
translation="MNCKKIFITSTLISLFLPGVSPFSDVIQEDSNPAGSVYISAKY
MPTASHFGKMSIKEDSKNTOTVFGLLKDDMGVKTSPDSSNTNSITFEKDYSPRENN
PFLGFAAGIYSGMNGPRIEFVSYETFDVKNLGGNYKNNAHMYCALDTAAQNSYNGAG
LTTVMVKNENLTNLSLNACYDMLDGPVSYVACAGIGTDLVSVINATNPKLSYQ
GKLGISYSINSEASIFIGGHRFVIGNEFKDIATLKIFTSKTGISNPGFASATLDVCH
FGIEIGGRFVF"
BASE COUNT      307 a      139 c      149 g      269 t
ORIGIN
Alignment Scores:
Pred. No.:      6.37e-26      Length:      864
Score:          383.50      Matches:      100
Percent Similarity: 50.17%      Conservative: 46
Best Local Similarity: 34.36%      Mismatches: 108
Query Match:      25.40%      Indels:      37
DB:              6      Gaps:      12

US-09-846-808-1 (1-284) x AX042305 (1-864)
Qy 13 ValMetLeuIleLeuPheThrProHisValHisSerLeuAlaSerValLeu----- 28
Db 37 ATATCATAGTGTCAATTTTACCTGGTGTGCTCTTTCTGTGATTAATACAGGAAGACAGC 96
Qy 29 AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
Db 97 ACCCAGCAGCAGTGTGTACATTAGCCGCAAAATACATGCCAATGCCATCATTTTGGT 156
Qy 49 LysLeuLeuIleGlyGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68
Db 157 AAAATGTCAATCAAGAAGATCAAAAAATACTCAACGGTATTGGTCTAAAAAAAGAT 216
Qy 69 LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsnPhe 83
Db 217 -----TGGGATGCGCTTAAACACCATCAGATTCCTAGCAATACTAATCTT 261
Qy 84 AsnIle-----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
Db 262 ACAATTTTACTGAAAGACATATCTTTTCAGATATGAAACAAATCCGTTTTAGTTTC 321
Qy 99 SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 322 GCTGGAGCAATTTGGGTACTCAATGAAT---GGACCAAGAATAAGAGTTCGAAGTATCCTAT 378
Qy 119 GluThrPheHisIleLysAsnGly-----TyrLysArgIleAspCysGluLysHis 136
Db 379 GAACTTTTGTATGTAATAAACCTAGTGGCAACTATAAAAC---AACGCACACATGTAC 435
Qy 137 PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
Db 436 TGTGCTTTAGATACAGCAGCAAAATAGCACTAATAGCGCAGGATTAACTACATCTGT 495
Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
Db 496 ATGGTAAAAACGAAATTTAAACAAATATATCATTAATGTTAAATCCGTTTATGATATC 555
Qy 177 -----AspGlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAsp 194
Db 556 ATGCTTGATGGAATACCA-----GTTCTCCATATGATGTCAGGATTGGCAGCTGAC 609
Qy 195 ThrIleAspPhePheLysSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
Db 610 TTAGTGTCAAGTAATTAATGCTACAAATCCTAAATTTATCTTATCAGGAAGCTAGGCATA 669
Qy 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
Db 670 AGTTACTCAATCAATCTGAAGCTTCTATCTTTATCGGTGGACATTTCCATAGAGTTATA 729
Qy 235 GlyLysLysPheGluLysIleProValAsnTyrProCysAspTyrProSerProThrPro 254
Db 730 GGTAAATGAATTTAA-----GATATTGCTACCTTTAAATA 765
```

```
Qy 255 ProAsnSerLysProHisValHisThrThr-----AlaLeuAlaMetLeuSerIleGly 272
Db 766 TTTACTTCAAAAACAGGAATCTTAATCTCGCTTGGCATCAGCAACACTTGATGTTGT 825
Qy 273 TyrTyrGlyGlySerIleGlyIleLysPheIle 283
Db 826 CACTTTGGTATAGAAATTCGAGGAGGTTTGTGA 858

RESULT 32
CRU50833      1265 bp      DNA      linear      BCT 14-JUL-1996
LOCUS      Cowdria ruminantium Highway major antigenic protein 1 (map1) gene,
DEFINITION      complete cds.
ACCESSION      U50833
VERSION      U50833.1 GI:1418255
KEYWORDS      .
SOURCE      heartwater rickettsia strain-Highway.
ORGANISM      Cowdria ruminantium
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Cowdria.
REFERENCE      1 (bases 1 to 1265)
AUTHORS      Reddy,G.R., Sulsona,C.R., Harrison,R.H., Mahan,S.M., Burrridge,M.J.
and Barbet,A.F.
TITLE      Sequence heterogeneity of the major antigenic protein 1 genes from
Cowdria ruminantium isolates from different geographical areas
JOURNAL      Clin. Diagn. Lab. Immunol. 3 (4), 417-422 (1996)
MEDLINE      96400830
REFERENCE      2 (bases 1 to 1265)
AUTHORS      Reddy,G.R.
DIRECT SUBMISSION
TITLE      Submitted (06-MAR-1996) Roman G. Reddy, Pathobiology, University of
JOURNAL      Florida, 471 Mowry Rd., Gainesville, FL 32611-0880, USA
FEATURES
source
1..1265
/organism="Cowdria ruminantium"
/strain="Highway"
/db_xref="taxon:779"
383..1246
/gene="map1"
383..1246
/gene="map1"
/feature="map1"
/note="surface protein"
/codon_start=1
/transl_table=11
/product="major antigenic protein 1"
/protein_id="AAC44144.1"
/db_xref="GI:1418256"
/translation="MNCKKIFITSTLISLFLPGVSPFSDVIQEDSNPAGSVYISAKY
MPTASHFGKMSIKEDSKNTOTVFGLLKDDMGVKTSPDSSNTNSITFEKDYSPRENN
PFLGFAAGIYSGMNGPRIEFVSYETFDVKNLGGNYKNNAHMYCALDTAAQNSYNGAG
LTTVMVKNENLTNLSLNACYDMLDGPVSYVACAGIGTDLVSVINATNPKLSYQ
GKLGISYSINSEASIFIGGHRFVIGNEFKDIATLKIFTSKTGISNPGFASATLDVCH
FGIEIGGRFVF"
BASE COUNT      444 a      175 c      194 g      452 t
ORIGIN
Alignment Scores:
Pred. No.:      1.01e-25      Length:      1265
Score:          383.50      Matches:      100
Percent Similarity: 50.17%      Conservative: 46
Best Local Similarity: 34.36%      Mismatches: 108
Query Match:      25.40%      Indels:      37
DB:              1      Gaps:      12

US-09-846-808-1 (1-284) x CRU50833 (1-1265)
Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeu----- 28
Db 419 ATATCATAGTGTCAATTTTACCTGGTGTGCTCTTTCTGTGATTAATACAGGAAGACAGC 478
Qy 29 AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
Db 479 AACCCAGCAGCAGTGTGTACATTAGCGCAAAATACATGCCAATGCCATCATTTTGGT 538
```



```

/db_xref="GI:13512588"
/translation="MNNKVLQNEVLLLIALLPSPYSLHNSIDIEYNEOPIKKLSRLNKNTI
KSQKLHNYKYLILKKIKISSIGNEWFLKFSNDYSNSQVIYYLRTLSSLKVSIALPDFK
NIFLNYKVLDTSLRNSFNSANYFESKHHNITNCFVSRAYFYLSCSY"
3882..4832
/gene="p30-17"
CDS
3882..4832
/gene="p30-17"
/note="P30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-17"
/protein_id="AAK28683.1"
/db_xref="GI:13512589"
/translation="MNVAKVFILCIIFLPSPSLFATNNNVFVHEIGKSIIGHFYIGVO
YKPTPHNRESIADDSFTNLLAISHTKDYLFYSSTEVRGILFSIPOEQNLHLYATGG
STTLTLKDSNKFPGNPTTYTDNLLGVGGIVGYSINMLRLEAEAFYERKNIAKPTGY
NYDEYFAIATVYKTKKAYVHYHCKMKNNTGILSSFLVNTCYDFTLTKAKIAPYLCL
GVGGDFIDFGLQTKKASYAKAGLSAISPNTLFFVDGSEHGMNNOFPGLLVDYPT
DISVMPSGDNATAYSEFTTMLAKLNMIFLAGSIGIRFIS"
4848..5699
/gene="p30-16"
CDS
4848..5699
/gene="p30-16"
/note="P30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-16"
/protein_id="AAK28684.1"
/db_xref="GI:13512590"
/translation="MYKLYLSFIIISLAQLFSGFAFSIDKNNNIHGSYITIKYOPTI
SNFKNFIKETDTEPDIGDIIAPNTNDFELKHNFVSLYHKDSKVFYENDLSGL
ALSTGLLVKLRNIEFEGSKYKFDTKRLAYHSREGHKFFAIPRTSNFGVIPNEQNTV
AKNGISILNIIKCSETKEKNETPICIGIGDFTFIEDVMRVKTYQKQGVISY
ITPKLYSISQYHKVIGNKFKFLPIQPVALKRINDSPEDKVTALLTLDLEHFSSE
IGLSFIF"
5716..6510
/gene="p30-15"
CDS
5716..6510
/gene="p30-15"
/note="P30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-15"
/protein_id="AAK28685.1"
/db_xref="GI:13512591"
/translation="MYNIINVIKTYIALAFLLIPRVSFSLIGNIEKSIKLSVHIN
SQYKPSISQISNYLIQENNSKEKKNILNLSNNTITYNMQLENSTNFRFTIGYFFKR
LRFAVEDSEFHKIDKSDSKANLSKYSKYMNEDFQNTIATDNKLSITSAIYNICY
DILINNTVLPHLCTAVGICSTGFNDMRFKLYQKIGLGLYLSNVMFLPFNVVYHK
VMRNKLNLTQYSVDINAFDAITVLANTDIGYFSGSEGVGRFIFN"
6525..7394
/gene="p30-14"
CDS
6525..7394
/gene="p30-14"
/note="P30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-14"
/protein_id="AAK28686.1"
/db_xref="GI:13512592"
/translation="MLQRLNFINILAFLLLPFPQSFYLIHDEHITQNVNGLYISSQ
YKPSIPYVSGGPRLIEGSHKTVELKLANDVTHVTVYVLKNDTKENTPYSAKFRNSLIN
LSGAPYVSGGPRLIEGSHKTVELKLANDVTHVTVYVLKNDTKENTPYSAKFRNSLIN
HSNMSYTFIKNNGISILSVNINGCYDIAFSNVKISPYVCAGIGDFTLETWHKIF
AYQKFGISILVSPSISIFANGHYHKVNDNVFNKLNHVKYIYKLODAPTITSARAKLRI
GYGSGEVGRFVF"
7419..8255
/gene="p30-13"
CDS
7419..8255
/gene="p30-13"
/note="P30 family member"
/codon_start=1

```

```

/transl_table=11
/product="major outer membrane protein P30-13"
/protein_id="AAK28687.1"
/db_xref="GI:13512593"
/translation="MNNKKSLLIGITILLSSFSSLPKAFSVINHSDISSNVNGLFTGT
QYRPVSHFSGFTVRENTIATQOLVSLNTKNKNHIIERTNFSGIYTKAFQDNAAFS
GAIGSYPEGLKFEIEISYKFGVKSTKNYQISNFIALARQTTSSNDKYSYYVMK
NIGLSVASVMINGCYNMSFYNLVSPYICAGIGEDFIEFDLYIKLAYOGKLGVNS
LSSRENIFADMYHKVIGNQPKNLNVAIHAVALDTPKVTSAITLNVAFEGVEGIRE
IL"
8266..8823
CDS
8266..8823
/codon_start=1
/transl_table=11
/product="unknown function U3"
/protein_id="AAK28688.1"
/db_xref="GI:13512594"
/translation="MATYMIHKAKIIEPFSILTYYIILSLNNVQDNHSEIICIT
KLYKILISIRKKLINVDNIADVILNKNCSKFKSTOSYEMFYFTKNIKDR
ILNFSLSYFIKYNVTQRKNSSEYKSTYKSHNFNATYTSFIHQHKLTYKYSITNOH
NKSTKSFNKKILAKNILNQNSEYE"
8816..9685
/gene="p30-12"
CDS
8816..9685
/gene="p30-12"
/note="P30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-12"
/protein_id="AAK28689.1"
/db_xref="GI:13512595"
/translation="MNSKTFEILSGILICLAALPIQSFSSSNVYTNKHTGLYIS
GLYKPSHFSDFSIKETYNTEALFGKQDIISSILRNKKTQYNNNVNVPYAKFQD
DPASIAVAVIANNPRIIEGSEEDVKNGNVTIDAHRYLALAREKTSYYLSS
KCKNKKYIILKNGISIVSIIINGCYDILSDSKVSPYICTGEGGDFIEFSAIREKFA
YQKIGISYSSNIIILFTDGYIHKVINSQFNLNVEHVNNELTDPKVTSAITAFPLNI
EYFGFGLKFI"
Alignment Scores:
Pred. No.: 4..4e-24 Length: 28254
Score: 383..50 Matches: 101
Percent Similarity: 52.05% Conservative: 51
Best Local Similarity: 34.59% Mismatches: 112
Query Match: 25.40% Indels: 28
DB: 1 Gaps: 8
US-09-846-808-1 (1-284) x AF078553 (1-28254)
QY 3 LysArgSerAsnArgLysPheValLeuTrpValMetLeuLeuPheThrProHisIle 22
||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 24810 AAGATGAATTACAAAAGATTGTTGAGGTGTACGCTGAGTACATTGTTTTCTTA 24751
QY 23 SerLeuAlaSerValLeuAsnAspHisAsn-----SerMetTyrVal 36
||| ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 24750 TCT---GATGGTGCTTTTCTTGATGCAAAATTTCTTGAAGGGAGAGACATTTATATA 24694
QY 37 GlyIleGlnTyrLysProAlaArgGlnHisLeuSerLysLeuLeuIleLysGluSerAla 56
||| ||| ||| ||| ||| :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 24693 GGTAGTCAGTATAAAGTTGGTATTCCCAATTTTAGTAATTTTTCAGCTGAAGAACAAATT 24634
QY 57 AlaAsn---ThrValGluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThr 75
||| ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 24633 CCTGGTATTACAAAAGATTTTGGTGTAGGTCTTCAT-----AAGTCT 24589
QY 76 GlyIleLysAspAsnThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArg 95
||| ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 24588 GAGATAAATACACACAGCAATTTTACACGATCATATGACCCCTACTTACGACGAGTCTTT 24529
QY 96 LeuGlyPheSerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGlu 115
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 24528 GCAGGGTTTAGTGGTATCATTTGATTAAT---CAGCTTAGGTAGATTTGAA 24472
QY 116 LeuSerTyrGluThrPhe-HisIleLysAsnAsnGlyTyrLysArgIle--AspCysGlu 134
||||| ||| ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

```


Db	745	TTTACTTCAACTAGTAAGCATCATCTATACCTTAAT---CCTGGCTTTGCATCAGCAACA	801
Qy	269	LeuSerIleGlyTyrTyrGly 275 ::: :	
Db	802	CTTGATGCTGCCATTTGCGT 822 :	
RESULT 35			
AF325176			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
gene			
CDS			
BASE COUNT	283 a	130 c	148 g
ORIGIN			
Alignment Scores:			
Pred. No.:	9,01e-26	Length:	816
Score:	381.50	Matches:	93
Percent Similarity:	54.92%	Conservative:	41
Best Local Similarity:	38.11%	Mismatches:	83
Query Match:	25.26%	Indels:	27
DB:	1	Gaps:	12
US-09-846-808-1 (1-284) x AF325176 (1-816)			
Qy	13	ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn	32
Db	19	ATATCATTTAGTGCATTTTACCTGTGTGTCATTTTCTGATGTAATACAGGAAGAAC	78
Qy	33	-----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer	48
Db	79	AATCCAGTAGTGTGTTTATATATAGCGCAAAATACATGCCAACTGCATCATTTTGGT	138
Qy	49	LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysAsp	68
Db	139	AAAATGTCAATTAAGAAGATTCTAGAGATCTAAAGCAGTATTTCCTTAAAAAAGAT	198
Qy	69	LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsn-----Phe	83
Db	199	-----TGGGATGGAGTTAAACACACCACCGGTACACCAATTCATTTT	243
Qy	84	AsnIleLys---TyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPheSerGlyIle	101
Db	244	ACTGAAAAAGACTATTCTTTCAAAATATGAAATAATCCATTCTTAGGTTTTCAGGAGCA	303
Qy	102	PheGlyTyrTyrTyrAsnAsnPheArgIleGluSerGluLeuLysHisPheThrPhe	121
Db	304	GTTGGTACTCAATGAAT---GGACCAAGAATAAGATTGAAGTATCTTATGAACATTC	360
Qy	122	HisIleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHisPheAlaLeu	139
Db	361	GACGTAAATAATCCAGCGGTAACTACAAAAC-----GATGCACATATGATTGT	411
Qy	140	AlaLysGluIleSerGlySerAsnAsnProAlaAsnLysTyrValThrLeuIle	159
Db	412	GCTAGATACAGCATCAAGCTCTACTGCAGGAGCACTACATCT---GTTATGGTAAAA	468
Qy	160	AsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal-----Asp	177
Db	469	AATGAAAAATTTAAGATATTTTCATTAATGCTAAACGCATGTTATGACATAATGCTTGAC	528
Qy	178	GlyLeuLysHisAsnIleThrTyrSerCysLeuGlyPheGlyValAspThrIleAsp	197
Db	529	GGAATGCCA-----GTTTCTCCATATGATGTCAGGCATGCTGACTAGTATCA	582
Qy	198	PheLeuSerLysTyrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThr	217
Db	583	GTAATTAACGCTCAAAATCTAAATTTATCTATCAAGGAAAAATAGGAATAAGTTATCA	642
Qy	218	ValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLys	237
Db	643	ATAAACCCGGAAGCTTCTATCTTTATTGGTGGCAGCTTCATAGAGTATAGTAAACGAA	702
Qy	238	PheGluLysIle 241 : :	
Db	703	TTTAAGATATT 714 : :	
RESULT 36			
AF368013			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
gene			
CDS			
BASE COUNT	283 a	130 c	148 g
ORIGIN			
Alignment Scores:			
Pred. No.:	9,01e-26	Length:	816
Score:	381.50	Matches:	93
Percent Similarity:	54.92%	Conservative:	41
Best Local Similarity:	38.11%	Mismatches:	83
Query Match:	25.26%	Indels:	27
DB:	1	Gaps:	12
US-09-846-808-1 (1-284) x AF325176 (1-816)			
Qy	13	ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeuAsnAspHisAsn	32
Db	19	ATATCATTTAGTGCATTTTACCTGTGTGTCATTTTCTGATGTAATACAGGAAGAAC	78
Qy	33	-----SerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer	48
Db	79	AATCCAGTAGTGTGTTTATATATAGCGCAAAATACATGCCAACTGCATCATTTTGGT	138
Qy	49	LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysAsp	68

KEYWORDS

SOURCE heartwater rickettsia.

ORGANISM Cowdria ruminantium

Bacteria; Proteobacteria; alpha subdivision; Rickettsiales; Rickettsiaceae; Ehrlichieae; Cowdria.

REFERENCE 1 (bases 1 to 825)

AUTHORS Allsopp,M.T.E.P., Dorfling,C.M., Maillard,J.C., Bensaïd,A., Haydon,D.T., van Heerden,H. and Allsopp,B.A.
Ehrlichia ruminantium Major Antigenic Protein Gene (map1) Variants Are Not Geographically Constrained and Show No Evidence of Having Evolved under Positive Selection Pressure

J. Clin. Microbiol. 39 (11), 4200-4203 (2001)

MEDLINE 21539003

PUBMED 11682561

REFERENCE 2 (bases 1 to 825)

AUTHORS Allsopp,M.T., Dorfling,C., Maillard,J.-C., Bensaïd,A., Van Heerden,H. and Allsopp,B.A.

Direct Submission

TITLE Submitted (05-APR-2001) Molecular Biology, Onderstepoort Veterinary

Institute, Old Soutpansberg Road, Onderstepoort, Gauteng 0110,

South Africa

FEATURES

source Location/Qualifiers

1..825

/organism="Cowdria ruminantium"

/isolate="Mar87/7"

/db_xref="taxon:779"

<1..>825

/codon_start=1

/transl_table=11

/product="major antigenic protein MAP1"

/protein_id="AAK98148.1"

/db_xref="GI:15429328"

/translation="FITSTLILSVLPFGVSFSDVIOEDSNPAGSVVISAKYMPYASH
FGKMSKEDSKNTQTVFGLKWDGVKTPSDSNTNSTFTTEKDYFRYENNPFLGFA
GAIGYSNMPRIEFESYETFDVKNPGKNNAHMYCALDRAQNSTNGAGLTTYSVM
VNNENLISMLNAGTIDMLDGIPVPCVACIGIDLVSVINATPKLISYQKLGIS
YSINSASIFIGGHRVIGNEFKDITATIKFTSKTGISNPGFASATLDVCHRGIEI"

BASE COUNT 293 a 139 c 138 g 255 t

ORIGIN

Alignment Scores:

Pred. No.: 1,71e-25 Length: 825
Score: 378.50 Matches: 91
Percent Similarity: 52.63% Conservative: 39
Best Local Similarity: 36.84% Mismatches: 90
Query Match: 25.07% Indels: 27
DB: 1 Gaps: 10

US-09-846-808-1 (1-284) x AF368008 (1-825)

QY 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeu----- 28
 Db 19 ATATCATAGTGTGTCATTTTACCTGGGTGTCCTTTTCTGATGTAATACAGGAAGACGC 78
 QY 29 AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
 Db 79 AACCCAGCAGCAGTGTTCATTAGCGCAAAATACATGCCAACTGCATCACATTTTGGT 138
 QY 49 LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68
 Db 139 AAATGTCATCAAGAAGATCAAAAATACTCAAAACAGTATTTTGTCTCAAAAAAGAT 198
 QY 69 LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAspPhe 83
 Db 199 -----TGGGATGGGTAAACACCATCATCATCTAGCATATAATCT 243
 QY 84 AsnIle-----LysTyrAsnProTyrTyrGluAsnAsnArg----LeuGlyPhe 98
 Db 244 ACAATTTTACTGAAAAGACTATTCTTTCAGATATGAAAACAATCGGTTTATAGGTTTC 303
 QY 99 SerGlyIlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
 Db 304 GCTGGGCAATGGGTACTCAATGAAT---GGACCAAGAATAGAGTTTCGAAGTATCCTAT 360

QY 119 GluThrPheHisIleLysAsnAsnGly-----TyrLysArgIleAspCysGluLysHis 136
 Db 361 GAAACTTTTGTATGTAATAAACCCAGCGGCAACTATAATAAAC---AACGCACACATGTAC 417
 QY 137 PheAlaLeuAlaLysGluIleSerAsnAsnProAlaAsnAsnLysTyrVal 156
 Db 418 TGTCTCTTAGATACAGCAGCAGCAAAATAGCACATATGCGCAGGATTAACTACATCTGTT 477
 QY 157 ThrLeuLeuAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspVal 176
 Db 478 ATGGTAAAAACGAAATTTAACAAATATATCATTAATGTTAAATGCGTGTATGATATC 537
 QY 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValAsp 194
 Db 538 ATGCTTGTGAATACCA-----GTTTCTCCATATGATGTCAGGTATTTGGCCTACAC 591
 QY 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAla 214
 Db 592 TTAGTGTGTCAGTAATTAATGCTACAAATCCTAAATATCTTATCAAGGAAGACTAGGCATA 651
 QY 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPhe 234
 Db 652 AGTTACTCAATCAATCTCTGAAGCTTCTATCTTATCGTGGACATTTCCATAGAGTTATA 711
 QY 235 GlyLysLysPheGluLysIle 241
 Db 712 GGTAAATGAATTTAAAGATATT 732

RESULT 40

AF368010

LOCUS

DEFINITION

MAP1 gene, partial cds.

ACCESSION

AF368010

VERSION

AF368010.1 GI:15429331

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..834

/organism="Cowdria ruminantium"

/isolate="Morgenswag2"

/specific_host="goat"

/db_xref="taxon:779"

/note="detected in goat blood from the farm Morgenswag,

Northern Cape, South Africa"

<1..>834

/codon_start=1

/transl_table=11

/product="major antigenic protein MAP1"

/protein_id="AAK98150.1"

/db_xref="GI:15429332"

/translation="FITSTLILSVLPFGVSFSDVIOEDSNPAGSVVISAKYMPYASH
FGKMSKEDSKNTQTVFGLKWDGVKVPSTENTNYSSLFTKDYSPRYENNPFLGFA

GAIGYSMNGPRIIEVSYETFDVKNPGGNYKNDAMHYCALDTAQOSATNGATLASSVY
IKNENLTNISMLNACYDIMLDGMPVSPYVCAGIGPDLVSVINATPKLSYOGKLGIS
YSINSGASIFIGGHFHRVIGNEFKDIATKIFTATNKVSTVANPGFASATLDVCHFGI
EI"

BASE COUNT 285 a 136 c 145 g 268 t
ORIGIN

Alignment Scores:

Pred. No.: 1.73e-25 Length: 834
Score: 378.50 Matches: 91
Percent Similarity: 53.04% Conservative: 40
Best Local Similarity: 36.84% Mismatches: 89
Query Match: 25.07% Indels: 27
DB: 1 Gaps: 10

US-09-846-808-1 (1-284) x AF368010 (1-834)

Qy 13 ValMetLeuIleLeuPheThrProHisIleSerLeuAlaSerValLeu----- 28
Db 19 ATTTCAATTAGTGTCAATTTTACCTGGCGTGTCTTTCTGTATGTAATACAGGAGGATAGC 78
Qy 29 AsnAspHisAsnSerMetTyrValGlyIleGlnTyrLysProAlaArgGlnHisLeuSer 48
Db 79 AACCAGCTGGTAGTGTTCATATAGCGCAAAATACATGCCAATCTTCACACITTTGGT 138
Qy 49 LysLeuLeuIleLysGluSerAlaAlaAsnThrValGluValPheGlyLeuLysLysAsp 68
Db 139 AAAATGTCATTAAGAAGATTCAAAAATATCTCAACACAGTGTGGTCTAAAAAAGAT 198
Qy 69 LeuLeuAsnAspLeuLeuThrGlyIleLys-----AspAsnThrAsnPheAsn 84
Db 199 -----TGGGATGGAGTTAAAGTAGTACCACATCACAAACACCACTACTCT 243
Qy 85 Ile-----LysTyrAsnProTyrTyrGluAsnAsnArg---LeuGlyPhe 98
Db 244 TCACITTTTACTGAAAAGATTATCTTCAGATATGAACAACATCCGTTTGTAGTTTT 303
Qy 99 SerGlyIlePheGlyTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyr 118
Db 304 GCTGGAGCAATTGGGTACTCAATGAAT---GGACCAAGAATAGAGTTCCGAAGTATCCTAT 360
Qy 119 GluThrPheHisIleLysAsnGly-----TyrLysArgIleAspCysGluLysHis 136
Db 361 GAACTTTTGTATGATAAACCAGGTGGCAGTACAAAAAT---GATGCACACATGTAC 417
Qy 137 PheAlaLeuAlaLysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrVal 156
Db 418 TGTGCTTTAGATACAGCACACAGAGTGTCTAATGGTGGCAACATTACTTCATCTGT 477
Qy 157 ThrLeuIleAsnAsnGlyIleSerLeuThrSerAlaLeuIleAsnValCystyrAspVal 176
Db 478 GTGATAAAAAATGAAAAATTAACAATAATATCATAATGTTAAATGCGTGTATGATATA 537
Qy 177 -----AspGlyLeuLysHisAsnIleIleThrTyrSerCysLeuGlyPheGlyValasp 194
Db 538 ATGCTGTATGGAATGCCA-----GTTCTCCATATGTATGTCAGGTATGTCCTGAT 591
Qy 195 ThrIleAspPheLeuSerLysTyrThrThrLysPheSerTyrGlnGlyLeuGlyAla 214
Db 592 TTAGTGTCAAGTAATTAATGCTACAATCTCAATATCTTATCAGGAAGCTAGGCATA 651
Qy 215 SerTyrThrValSerProGlnValSerValPheIleGluGlyTyrThrHisGlyLeuPhe 234
Db 652 AGTACTCAATCAATCTCGAGCTTCTATCTTTATCGGTGGACATTTCATAGAGTTATA 711
Qy 235 GlyLysLysPheGluLysIle 241
Db 712 GGAATGAATTTAAAGATATT 732

Search completed: October 19, 2002, 04:29:56
Job time : 1921 secs

This Page Blank (uspto)

ORGANISM Ehrlichia chaffeensis
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
REFERENCE Yu.X.J., McBride, J.W., Zhang, X. and Walker, D.H.
AUTHORS Characterization of the complete transcriptionally active Ehrlichia
TITLE chaffeensis 28 kDa outer membrane protein multigene family
JOURNAL Gene 248 (1-2), 29-68 (2000)
REFERENCE 2 (bases 1 to 14759)
AUTHORS Yu.X.J., McBride, J.W., Zhang, X. and Walker, D.H.
TITLE Direct Submission
JOURNAL Submitted (01-FEB-2000) Pathology, University of Texas Medical
Branch, 301 Univ. Blvd., Galveston, TX 77555-0609, USA
FEATURES
Source
1. 14759
/organism="Ehrlichia chaffeensis"
/strain="Arkansas"
/db_xref="taxon:945"
400..1731
/gene="clpX"
400..1731
/gene="clpX"
/note="ATP-dependent protease"
/codon_start=1
/transl_table=11
/product="ClpX"
/protein_id="AAF73407.1"
/db_xref="GI:8132824"
/translation="MADNEKSCSFCGKIHSVRKLIAGSVFTCEIDLCSGIL
QESRSYKTDKLRKPKIKKVLDEYVIGHEHKKVLSVAVNYHKKLSLVISEV
EIKSNVLLIGTGSGKTLARTLARIQIPAMADATLLEAGVGVGDVENILKLL
QAANENVAQRGIIYIDEVOKISRSKNTSTTRDVSGEGVOQALLVIEGVTSVPP
QGRKHPHQAESLHKSVMQLSNRKVSSTTKVTALYIEIIQDFEQLSDQEEV
LELYEGATRDVDKQYKEDNYFALCKDKTESTDNGNHNVMKNEGFRVLSL
TNACYDMIVSNLSVPSACIGIGGITNFGGTNIHLFKAKLGLGLSPKTVIFA
NGYVKTDMNSVNSLVSQIPLELKEAPKHIDFIACFNADNYGGEVGLREIL"
2336..3051
/gene="p28-1"
2336..3051
/gene="p28-1"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-1"
/protein_id="AAF73408.1"
/db_xref="GI:8132825"
/translation="MLIFTPHISLASVLNDHNSMYGVIOYKPAQOHLKLIKESAA
NVEVEGLKDLNLLTGKDNFNINRYENNNRNLGFGSGIFGYNNKRIESE
LSYETPHKNNGYKRIDCEKHFALAKEISGGNNPANNKYVTLINGISLTSALINVC
YVDGLKHLNITYSCLGFGVDITDLSKYTKFYSYQGLGASVTSPOVSVEIGYIH
GLGKFEKIPVNYPCDYSPTPPNSKPHVHTTALAMLSIGYVGSIGIKFIL"
3158..3604
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAF73409.1"
/db_xref="GI:8132826"
/translation="MNNKSRGQVLLFAAYSIFKYINSCTSKPLRSVKCFVYNPLIYF
YKSTFLGKFFVNYNSKTSNLYHRSFLKFLGFIITNFKISYLNFTNPRVLIY
YKILKNKNSSINYLKNTYINPTATSFINSVIIGIYPGFVNYN"
3614..4507
/gene="p28-2"
3614..4507
/gene="p28-2"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-2"
/protein_id="AAF73410.1"

/db_xref="GI:8132827"
/translation="MSYAKVFILICILILVPSLSFAIVNDFLKDNIHGYIGGOYKP
GVPRFKRLVTNNRELMSDEECSTLPHVQSVQAGTLPPEALEELADGKPEPGY
LYETIPNYTNNKLLGAGVIGYSTHERVEAYEAFNLTAPAGYLHKNEVEFA
LATMTDKHPQSAEDKYIYMNKTITLSPFIINACYDFILKKTNRNVPATLCLGVGN
FIDFLQVSKFAYQAKVIGSYFVSFNIAFFIDGSPHGLNNOFSFSPVVDYSSGFP
TISAKNANPLTSSIGIRFS"
4524..5381
/gene="p28-3"
4524..5381
/gene="p28-3"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-3"
/protein_id="AAF73411.1"
/db_xref="GI:8132828"
/translation="MOKLYISFIILSGLLLPKPVFCHMNNNDGSYVYIKYQTTTH
FKNFYIKETDFTQEPILGAKITANTKTFTDLKENFSPLHQTSYKSYQNDLIGL
SVGLFVKSFRIEFGAYKNFTKLARYKSKDKYKFAIPKRSKHGFDJNTFGYTVAK
NNGIISININLCSETKYSFTPYICIGYGGDFLEIFDVMRIKFAVQGVGYSPIT
SKLILSINGQYHKVIGNKFEPLLPVQPVELARLVNTKTSKIDIDQDVATSLNLEHFS
SEIGLSFIF"
5394..6212
/gene="p28-4"
5394..6212
/gene="p28-4"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-4"
/protein_id="AAF73412.1"
/db_xref="GI:8132829"
/translation="MYMKNKHVCYIVTVTLFLLPIELSLALIGNEVDLKVSS
TVVSQYKPSIFHFRNFQIESHPKSKSESEKIKANLNILKSNAINLQFODNTSF
SGTIGYFKGLRLAEAGCYQEFNVKSNNSLISSNKYHSRIHDENYAITNNKUSIA
SIMVNTCYDISINNTSIVPYLCTGIGEDLVGLFNTIFPKLAYQGVKSYSLNNILL
PSEDIYHKVMGNRFKNLYMQVYVADPNISEETIPIKALDIDYFGSGEIGREMFN"
6228..7115
/gene="p28-5"
6228..7115
/gene="p28-5"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-5"
/protein_id="AAF73413.1"
/db_xref="GI:8132830"
/translation="MTKKNFVNVNLTFLLEPLKSFYANNNTITOKVGLYISGQ
YKPSIPHKNSVEKNDKVDLIGLTDVYITEHLIRONTKENTHYIAKKNNEINF
SSAIGYSGQSPRLIESYGDVNVKNYAVQDVNRYFALVREKNGSNFSKPHEH
SOPSDNPKKSFYTLKNGVFAVSIINGCYDFSNFTTISPYICIGYGGDFIEFFE
VMHIFACOSKVGISYPIPSITIFADAHYHKVNNKNNLHVKYSLKNSPTTISA
TAKLNEYFGSEVGMRFIF"
7137..7976
/gene="p28-6"
7137..7976
/gene="p28-6"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-6"
/protein_id="AAF73414.1"
/db_xref="GI:8132831"
/translation="MSKKKFTITGTVLASLLSFLSIESFSAIHNHNTGNTSGIYITG
QYRPGVSHFSNFKETNVDITQLGVKKKSSSIDPNTYINFGQPYVTVPQDAAFSF
GAIGYSPELSRELEGSYKFDKDPKADAFRRFALARTSTVYDQAKYTM
KNNGLSVASIMINGCYDLSFNILVSPYICAGIGEDIEFFDTHLHKIAYOGKLGISY
YFFPKINVFAGGYHRYIGNKFNKLNHNVTVPDEFPKATSAVATLNVAYFGGAGVK

This Page Blank (uspto)


```
CDS
8214..8591
/note="Orf378"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAF73415.1"
/db_xref="GI:8132832"
/translation="MNCYLYGKPLNQIFYGIFSRNFQNNLTIIIPNDSKCGFYTL
WNPALHYTYLTGSSYRNPFDDLYENIIOCKLLNYSRVNLQHNKNTLVIPIPN
AREFSNEIRVNISINKESSYEC"
8581..9432
/gene="p28-7"
8581..9432
/gene="p28-7"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-7"
/protein_id="AAF73416.1"
/db_xref="GI:8132833"
/translation="MSAKKLFITIGSLVCLVLYLPTKSLSNLNNNTKCTGLYVS
QYKPVHFSNFSKLTETDYDKELGLKANDIKSIDITTNKFNPIPIYKFKODNAVS
FSAAGYISODSPREVEWSYEEDVKFNQYVSEAFRIALARGIDNQLQYPTNK
YVYKNGLSVASIIIGCYDFSLNKLKVPYICVGGDIIIEFFSAVSFKAYQGV
GISYPLFSNMIIFADGYHKVIGNKNLNLVQHVSLNSHPKSTFAVATLNVEYFGSE
FGLKFTF"
9441..10268
/gene="p28-8"
9441..10268
/gene="p28-8"
/note="component of 28-kDa outer membrane protein
multigene family"
/codon_start=1
/transl_table=11
/product="p28-8"
/protein_id="AAF73417.1"
/db_xref="GI:8132834"
/translation="MSKKNFITIGATLIHMLPNISFPETINNNTDKLSGLYSQYK
PGISHPSKSVKELYNDIOLIGLHNAISTLNINLDFNPIYKVFQNNNTSFGKA
IGYSDPTGARFELGSEYEDFTDGDCLIKDTIRYFALARNPSSPSNNYVWRN
DGSVTFIFNGCIDYFLKDLVSPYCVGVGGDFIEFFDHALIKLAYQKLGINYHL
STQASVFDIGYHKVIGNKNLNLVQHVASTDFGPVYAVATLNIGYFGEIGRLTF"
10278..11135
/gene="p28-9"
10278..11135
```

Alignment Scores:

```
Pred. No.: 2,6e-126 Length: 14759
Score: 1510.00 Matches: 284
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
```

US-09-846-808-1 (1-284) x AF230642 (1-14759)

```
QY 1 MetSerLysArgSerAsnArgLysPheValLeuTrpValMetLeuLeuLeuPheThrPro 20
|||||
Db 2197 ATGAGTAAAGGAGTACAGAAAGTTGTTTGTGGTGTATGTAATATATTATACACCA 2256
|||||
QY 21 HisLeuSerLeuAlaSerValLeuAsnAspHisAsnSerMetTyrValGlyLeuGlnTyr 40
|||||
Db 2257 CATATTTCTTTAGCATCAGTTTAAATGACCAATATCTATGATGTTGTTTTCAGTAC 2316
|||||
QY 41 LysProLahArgGlnHisLeuSerLysLeuLeuLeuLysGluSerAlaLaAsnThrVal 60
|||||
Db 2317 AAGCCAGCCAGACCAACCTATCAAAACTCTTATAAAGAAAGCGCAGCAACACTGTA 2376
|||||
QY 61 GluValPheGlyLeuLysLysAspLeuLeuAsnAspLeuLeuThrGlyLysAspAsn 80
|||||
Db 2377 GAAGTTTTCGGGTAAAGAAAGACTTACTAAATGATCTATTAACTGGTATTAAGATAAT 2436
|||||
```

```
QY 81 ThrAsnPheAsnIleLysTyrAsnProTyrTyrGluAsnAsnArgLeuGlyPheSerGly 100
|||||
Db 2437 ACTAATTTTAAATAAATACACCCATATATGAAATAATATCGTTTAGGTTTTCCGGT 2496
|||||
QY 101 IlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThr 120
|||||
Db 2497 ATATTTGGATATTACTATAACAAAAATTCAGATAAATCTGAACCTTCTTACGAAAT 2556
|||||
QY 121 PheHisIleLysAsnAsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAla 140
|||||
Db 2557 TTCACATAAAGAACACACGGATATAAAGAAATGATTGTGAAAGCATTTTGTCTTTGCA 2616
|||||
QY 141 LysGluIleSerGlyGlySerAsnProAlaAsnAsnLysTyrValThrLeuIleAsn 160
|||||
Db 2617 AAGGAATAAGCGGTGGTAGTACATCTCCCAATATAAATAATATGTTACCTTAATAAT 2676
|||||
QY 161 AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCysTyrAspValaspGlyLeuLys 180
|||||
Db 2677 AATGGTATTTCACTTACCTCAGCTCTGATCAAGCTTTGTTATGATGTTGACGGACTTAA 2736
|||||
QY 181 HisAsnIleIleThrTyrSerCysLeuGlyPheGlyValaspThrIleAspPheLeuSer 200
|||||
Db 2737 CATATAATAATAACGTATTTCATGTTTAGGATTTGGAGTAGACACAAATGATTTTCTCAGT 2796
|||||
QY 201 LysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerPro 220
|||||
Db 2797 AGTATATACAAAGTTTTCATACCAAGAAATTAGAGCTAGTTATCTGTTCTTCCT 2856
|||||
QY 221 GlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLys 240
|||||
Db 2857 CAAGTATCGGTTTTTATAGAAGCTTATTATCATGTTTATTGTTAAATAATTGACAAA 2916
|||||
QY 241 IleProValAsnTyrProCysAspTyrProSerProThrProProAsnSerLysProHis 260
|||||
Db 2917 ATTCTCTGTAATTTATCCATGCTACTATCCATCCACCAACACCAACACCAAGCCACAT 2976
|||||
QY 261 ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyLe 280
|||||
Db 2977 GTTCACATACACACATTAGTATGTTAAGTATTGGATATTACGGTGGGAGTATTGGANTA 3036
|||||
QY 281 LysPheIleLeu 284
|||||
Db 3037 AAGTTTATATTA 3048
|||||
RESULT 2
ECU72291
LOCUS
DEFINITION
Ehrlichia chaffeensis strain Arkansas major outer membrane protein
Omp-1 multigene cluster, complete sequence.
ACCESSION
U72291 AF021338
VERSION
U72291.2 GI:13511827
KEYWORDS
SOURCE
Ehrlichia chaffeensis.
ORGANISM
Ehrlichia chaffeensis
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
1 (bases 14844 to 21136; 21479 to 22234),
Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y.
Immunodominant major outer membrane proteins of Ehrlichia
chaffeensis are encoded by a polymorphic multigene family
Infect. Immun. 66 (1), 132-139 (1998)
REFERENCE
2 (bases 1 to 27190)
PUBMED
9423849
JOURNAL
MEDLINE
98084465
PUBMED
9423849
AUTHORS
Ohashi,N., Rikihisa,Y. and Unver,A.
TITLE
Analysis of Transcriptionally Active Gene Clusters of Major Outer
Membrane Protein Multigene Family in Ehrlichia canis and E.
chaffeensis
JOURNAL
Infect. Immun. 69 (4), 2083-2091 (2001)
MEDLINE
21135566
PUBMED
11254561
REFERENCE
3 (bases 1 to 27190)
```

This Page Blank (uspto)

AUTHORS
TITLE
JOURNAL

Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y.
Direct Submission
Submitted (24-SEP-1996) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA

REFERENCE
4 (bases 1 to 27190)
Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y.
Direct Submission
Submitted (29-AUG-1997) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA

REFERENCE
5 (bases 1 to 27190)
Ohashi,N., Rikihisa,Y. and Unver,A.
Direct Submission
Submitted (29-NOV-2000) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA

REMARK
COMMENT
Sequence update by submitter
On Apr 2, 2001 this sequence version replaced gi:2853584 gi:2853273.

FEATURES
source
Location/Qualifiers
1. 27190
/organism="Ehrlichia chaffeensis"
/strain="Arkansas"
/db_xref="taxon:945"
315..956
/codon_start=1
/transl_table=1
/product="hypothetical transcriptional regulator"
/protein_id="AAC02808.2"
/db_xref="GI:13511828"
/translation="MSTHAKNKINKMETDIEKAHDLPIAGLIDKNNKKNFTNKGYN DDHSTKNTKEGVSTKDKNSTTKIRPHVPDECVGKEIKRQRMKMSQNLKGLG ITFOOVOKYEGTNRIVISRLYOLASVNEVDIMLKLOEDLNKISCDHTVASTSL KDNKEKRIPEHDSKIDSKEVLMVMRAYTCLKNEKVRNIYNLKVALSIDNQ"

gene
CDS
1136..2032
/gene="omp-1m"
1136..2032
/gene="omp-1m"
/note="omp-1 family member"
/codon_start=1
/transl_table=1
/product="major outer membrane protein OMP-1m"
/protein_id="AAK28658.1"
/db_xref="GI:13511829"
/translation="MILINKFLVHLNIMTKFSAIGIVLSLTLFACNVFASPIPID ESNESETAGYASHYNIQPRFSPISVKYKTDETEKELTFLKESTDTPEFKQS EFNKKGISPIIYRNITGFGAIGYSGGLRVELEGAFTDFVDKQKYYKKNRYFAL CKKDSIESTNSNGNHYVMKNEGFRVLSLTNACYDMIVSNSLVPACIGIGGITH FLGTNIHTLTKALGLGLFISPKTVIFANGYVYVTKDNSFTNLVQVPLELKAEPKH IDPIACFNADNYGGEVGLRFL"

gene
CDS
2498..3352
/gene="omp-1n"
2498..3352
/gene="omp-1n"
/note="omp-1 family member"
/codon_start=1
/transl_table=1
/product="major outer membrane protein OMP-1n"
/protein_id="AAK28659.1"
/db_xref="GI:13511830"
/translation="MSRSNRKFLVWMLILFTPHISLAVLNDHNSMYGVIOYKPAR QHLSKLLIKESAATVBFGLKDLNLLTGIDNTFNKYNPYVENNRGLGSGIF GYYNKNRLESLSYTFHTKNGYKRIDCEKHFALAKEISGSSNPNKYYVTLLN NGISLTALINVCVDGLKHNITYSCLGFGVDTIDFLSKYTKFSYQKLGASIV SPQVSFIEGTHGLGFKKFEKIPVNPYPCDYPSPTPPNKPHVHTTALAMLSIGYGG SIGIKFIL"

gene
CDS
3459..3905
/codon_start=1
/transl_table=1
/product="unknown function UN2"
/db_xref="taxon:945043"

translation="MNKRSRGVLLFAAYSIFKYINSCTSKFLRSVKCVNPLKYF YKSTFLKLGKFEVNYNYSKTSNLSLFLKFLGIIIRTNFKISYLNFTPRVLIY YKILKNKCNNSINLLNKYINPTATSFINSVIIGYPGFVNYN"
3915..4808
/gene="omp-1q"
3915..4808
/gene="omp-1q"
/note="omp-1 family member"
/codon_start=1
/transl_table=1
/product="major outer membrane protein OMP-1q"
/protein_id="AAK28661.1"
/db_xref="GI:13511832"
/translation="MSYAKVFIILICLLVPSLFAIVNNDFLKDNIHGFYIGQYKP GVPFRNRLVTNNIRELMSSDECRSTIPHMVQSVQAQTLPPPEALFELAKLHGKY LFTLLPVNFTYKNNLLCAGGIVGYSTHFRVEAEAFYEKENLTPAGYLVHKNFVEFA LATMDTKHPQSAEDKYYKMYKNTGTLSPFIILNACVDFILKTRNVPVLCIGVGN FIDFLQVSEKFAQKVGISYFVSPNIAFFIDGSHGLNNOFSDLPVVDYSSGGP TISAKFNANFLTSSIGIRFIS"
4825..5682
/gene="omp-1p"
4825..5682
/gene="omp-1p"
/note="omp-1 family member"
/codon_start=1
/transl_table=1
/product="major outer membrane protein OMP-1p"
/protein_id="AAK28662.1"
/db_xref="GI:13511833"
/translation="MOKLYIFSIILSGLLLPKVFCHQNNIDGSVVTIKYQITTPH FKFYIKETDFDQEPFGLAKITANTKFDLTKEFSPHQSDSYAQNDLIGLGL SVGLTSFIRIEFCGAYKNFTLARYKSKDGYKYFAIPRKSEHGLDNTGYTAK NNGISITSINILCSETKYKATPYICIGVGGDFIEFDVARIKFAVQGVGYSPIT SKILSLINGOYHKVIGNKPELLPYQVPELVKRLVTNKTSDIDODVTASLTNLNHS SEIGLSFIF"
5695..6513
/gene="omp-1t"
5695..6513
/gene="omp-1t"
/note="omp-1 family member"
/codon_start=1
/transl_table=1
/product="major outer membrane protein OMP-1t"
/protein_id="AAK28663.1"
/db_xref="GI:13511834"
/translation="MYMYNKKHYCYIVTVITLFFLLPTIESLSALIGNVEKDLKVS TYSSQYKPSIFHFRNFSIQESHPKSSSEEFKIKANLNILSNAINLQPDNTTSF SGTYGFSKGLRLEAGCYOEFNKVNNSNLISSNKYHSRIHNTAITNNKLSTIA SIMVTCYDLSINNTSIVPYLCTGIGEDLVGLFNTIHKLAYQGVKMGVSLINNILL FSDIYHKVNGNRKFNLYMOYVADPNISEETIPIKLAKLDIGYFGSEIGIRFMEN"
6529..7416
/gene="omp-1u"
6529..7416
/gene="omp-1u"
/note="omp-1 family member"
/codon_start=1
/transl_table=1
/product="major outer membrane protein OMP-1u"
/protein_id="AAK28664.1"
/db_xref="GI:13511835"
/translation="MTKKNFNVNVLFTLFLPKLSFTTYANNNTITQKVGLYISQO YKPSIFPKNFSEENDKVVDLIGLTDVITYITEHLRDNTEKNTYIAKFNKNFINF SSAIYVSGQPLETESYSGDFVNVKNYAVQDVNRYFALVREKNGSNFSPKPHET SOPSDSNPKFSFTLAKNNGVFAVSLINGCYDFSENNTTISPYVCIGVGGDFLEFE VMHKFACQSKVGISIPISITIFADAHKVINNKFNHLHVKYSIELANSPTTISA TAKLNIEYFGGEVGMRFIF"
7438..8277
/gene="omp-1v"
7438..8277
/gene="omp-1v"
/note="omp-1 family member"
/codon_start=1
/transl_table=1

This Page Blank (uspto)

product="major outer membrane protein OMP-1V"
/protein_id="AAK28665.1"
/db_xref="GI:13511836"
/translation="MSKKRTTIGTVLALLSLFSLIESFSAIHNHNTGNTSGIYITG
QYRGVSHFNFSVKEINVDITQLGVYKKSASSIDPNTYINFGQPYPTVTFDDNAASF
GAIGYSPELSRUELEGSYEKFDKDYSAKDAFFALARNSTVTPDAQKYIVM
KNGLSVASIMINGCYDLSFNNLVSPYICAGIGEDIEFFDLHLIKLAYOGKLGISY
YFFPKINVAGGYHVRVIGNKFNLVNHNVTLDFFPKATSAVATLNVAFEGGAGVK
FTF"
8300..8890
/codon_start=1
/transl_table=11
/product="unknown function UN3"
/protein_id="AAK28666.1"
/db_xref="GI:13511837"
/translation="MIYKELTRVGEVILAYLSFILSTYIFLVLNIRVNSLAICVI
SLRKNFNVSTKLKDKCRDYSNMNLYGKPLNQIFGYGIFSFIRNFQNTLI
IPDSKRGFTYLDNDPALHYTTLTGSEYNEFDILYENIIOCKKILLINRSVLNQ
HNKNTLVIIPIPNAREFSNIRVNSINKSEVSEC"
8880..9731
/gene="omp-1w"
8880..9731
/gene="omp-1w"
/note="omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1W"
/protein_id="AAK28667.1"

Alignment Scores:
Pred. No.: 5,45e-126 Length: 27190
Score: 1510.00 Matches: 284
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-846-808-1 (1-284) x ECU72291 (1-27190)

Qy 1 MetSerLysArgSerAsnArgLysPheValLeuThrValMetLeuLeuLeuPheThrPro 20
Db 2498 ATAGTAAAGAGGATACAGAAAGTTTGTGGGTATATGTAATATTTATACACCA 2557
Qy 21 HistSerLeuAlaSerValLeuAsnAspHisAsnSerMetTyrValGlyIleGlnTyr 40
Db 2558 CATATTTCTTAGCATCAGTTTAAATGACCAATATCTATGATGTTGGTATTCAGTAC 2617
Qy 41 LysProAlaArgGlnHisLeuSerLysLeuLeuLeuLysGluSerAlaAlaAsnThrVal 60
Db 2618 AAGCCAGCCAGACACACCTATCAAACTCTTTATAAAGAAAGCGCAGCAACACGTGA 2677
Qy 61 GluValPheGlyLeuLysLysAspLeuAsnAspLeuLeuThrGlyIleLysAspAsn 80
Db 2678 GAAGTTTTGGGTTAAAAAAGACTTACTAAATGATCTATTAAGTGGTATTAAGATAAT 2737
Qy 81 ThrAsnPheAsnLysLysAsnProTyrTyrGluAsnAsnArgLeuGlyPheSerGly 100
Db 2738 ACTAATTTTAAATAAATACACCAATATATGAAATATATCGTTAGGTTTTCCGGT 2797
Qy 101 IlePheGlyTyrTyrTyrAsnLysAsnPheArgIleGluSerGluLeuSerTyrGluThr 120
Db 2798 ATATTGGATATTACTATACAAAATTTTCAGATAGAACTGAATCTGAATCTTACGAAC 2857
Qy 121 PheHisIleLysAsnAsnGlyTyrLysArgIleAspCysGluLysHisPheAlaLeuAla 140
Db 2858 TTCACATAAAGAACACCGGATATAAAGAAATGATTGTGAAAAGCAATTTGCTCTTGA 2917
Qy 141 LysGluIleSerGlyGlySerAsnAsnProAlaAsnAsnLysTyrValThrLeuLeuAsn 160
Db 2918 AAGAAATAAGCGGTGGTAGTACAAATCCGCAATATAAATATGTTACCCCTAATAAT 2977
Qy 161 AsnGlyIleSerLeuThrSerAlaLeuIleAsnValCystTyrAspValAspGlyLeuLys 180

Db 2978 AATGGTATTTTCACTTACCTCAGCTCTGATCAACAGTCTTTTATGATGTTTCAGCGACTTAA 3037
Qy 181 HisAsnIleLeuThrTyrSerCysLeuGlyPheGlyValAspThrIleAspPheLeuSer 200
Db 3038 CATAATAATAACGATATTCATGTTTAGGATTTGGAGTAGACACAATGATTCTCTCAGT 3097
Qy 201 LysTyrThrThrLysPheSerTyrGlnGlyLysLeuGlyAlaSerTyrThrValSerPro 220
Db 3098 AAGTATCTACAAAGTTTTCATACCAAGGAAATTAGCAGCTAGTATATCTGTTCTCTCT 3157
Qy 221 GlnValSerValPheIleGluGlyTyrTyrHisGlyLeuPheGlyLysLysPheGluLys 240
Db 3158 CAAGTATCGGTTTTTATAGAAGTTATATCATGCTGTTTATTTGGTAAAAAATTTGAGAAA 3217
Qy 241 IleProValAsnTyrProCysAspTyrProSerProThrProAsnSerLysProHis 260
Db 3218 ATTCCTGTAATATTCATGCTGACTCATCCATCACCACCAACAGCAAGCCACAT 3277
Qy 261 ValHisThrThrAlaLeuAlaMetLeuSerIleGlyTyrTyrGlyGlySerIleGlyIle 280
Db 3278 GTTCACACTACAGCATTTAGTATGTTAAGTATTCGATATTCAGGTGGGAGTATTGGAATA 3337
Qy 281 LysPheIleLeu 284
Db 3338 AAGTTTATATA 3349
RESULT 3
AF078553 28254 bp DNA linear BCT 02-APR-2001
LOCUS
DEFINITION
Ehrlichia canis major outer membrane protein P30 multigene cluster
1, complete sequence
ACCESSION AF078553 AF078554 AF078555 AH006958
VERSION
KEYWORDS AF078553.2 GI:13512584
SOURCE
ORGANISM Ehrlichia canis.
Bacteria: Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Ehrlichieae; Ehrlichia; canis group.
REFERENCE
AUTHORS
TITLE
1 (bases 1 to 28254)
Ohashi, N., Unver, A., Zhi, N. and Rikihisa, Y.
Cloning and characterization of multigenes encoding the
immunodominant 30-kilodalton major outer membrane proteins of
Ehrlichia canis and application of the recombinant protein for
serodiagnosis
J. Clin. Microbiol. 36 (9), 2671-2680 (1998)
98371112
REFERENCE
PUBMED 9705412
AUTHORS
TITLE
2 (bases 1 to 28254)
Ohashi, N., Rikihisa, Y. and Unver, A.
Analysis of transcriptionally active gene clusters of major outer
membrane protein multigene family in Ehrlichia canis and E.
chaffeensis
Infect. Immun. 69 (4), 2083-2091 (2001)
JOURNAL
MEDLINE 21153566
PUBMED 11254561
REFERENCE
AUTHORS
TITLE
3 (bases 1 to 28254)
Ohashi, N., Unver, A., Zhi, N. and Rikihisa, Y.
Direct Submission
Submitted (16-JUL-1998) Department of Veterinary Biosciences, The
Ohio State University, 1925 Coffey Road, Columbus, OH 43210, USA
JOURNAL
REFERENCE
AUTHORS
TITLE
4 (bases 1 to 28254)
Ohashi, N., Rikihisa, Y. and Unver, A.
Direct Submission
Submitted (29-NOV-2000) Department of Veterinary Biosciences, The
Ohio State University, 1925 Coffey Road, Columbus, OH 43210, USA
JOURNAL
REMARK
COMMENT
Sequence update by submitter
On Apr 2, 2001 this sequence version replaced gi:3790556
gi:3790558 gi:3790557
FEATURES
source
1..28254
/organism="Ehrlichia canis"
/strain="Oklahoma"
/db_xref="taxon:944"

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)